Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
HUMAN RESOURCES RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *Stanford University*
Donald B. Lindsley, *Northwestern University*
Neal E. Miller, *Yale University*
Edwin B. Newman, *Harvard University*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Margaret K. Harlow, Managing Editor

CONTENTS

Superthreshold Reaction Potential (${}_sE_n$) as a Function of Experimental Extinction (n): H. G. YAMAGUCHI 391

Visual Duration Threshold as a Function of Word-Probability: D. H. HOWES AND R. L. SOLOMON 401

The Effect of Negative Incentives in Serial Learning: VI. Response Repetition as a Function of an Isolated Electric Shock Punishment: G. R. STONE AND N. WALTER 411

Pain Measurement by the Radiant Heat Method: Individual Differences in Pain Sensitivity, the Effects of Skin Temperature, and Stimulus Duration: J. E. BIRREN, R. C. CASPERSON, AND J. BOTWINICK 419

A Comparison of the Galvanic Skin Responses of Normals and Psychotics: A. S. PAINTAL 425

Studies of the Effect of Change of Drive: I. From Hunger to Thirst in a T-Maze: H. H. KENDLER AND S. LEVINE 429

Delay of Reward and Human Verbal Learning: I. J. SALTZMAN 437

The Rate of Learning a Tone—No-Tone Discrimination as a Function of the Tone Duration at the Time of the Choicepoint Response: M. U. ENINGER 440

The Role of Irrelevant Drive Stimuli in Learning Theory: M. U. ENINGER 446

The Effect of Training Procedures on the Relative Strength of Place and Direction Dispositions: F. R. FOSMIRE AND W. L. BROWN 450

Use of Spectral Hue-Invariant Loci for the Specification of White Stimuli: DOROTHEA JAMESON AND L. M. HURVITCH 455

American Psychological Association

Vol. 41 No. 6 June 1951

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc. The subscription rate per volume is $7.00, or $14.00 annually. Single copies are $1.25. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., Prince and Lemon Sts., Lancaster, Pa., or 1515 Massachusetts Avenue, N.W., Washington 5, D. C.

This JOURNAL publishes articles which are intended to contribute toward the development of psychology as an experimental science. Experimental work with normal human subjects is favored over work with abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental psychology or for behavior theory. Normally, articles of a length exceeding 25 printed pages cannot be accepted. However, a single experimental study or an integrated series of studies accomplished simultaneously must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors who request them are supplied with 50 free reprints without covers. All of the cost of the author's alterations in galley proof is charged to the author, and one-half of the cost of tables, cuts, and formulas.

Immediate publication can be given to articles whose authors assume the full cost of publication. Such articles must undergo the usual review by the Consulting Editors of the JOURNAL. No free reprints are given in the case of these articles.

Address all articles submitted for regular publication or for immediate publication to the editor, Arthur W. Melton, Human Resources Research Center, Lackland Air Force Base, San Antonio, Texas.

The JOURNAL adheres to the conventions concerning bibliographies, citations, preparation of tables and cuts, etc., as described by J. E. Anderson and W. L. Valentine, The preparation of articles for publication in the journals of the American Psychological Association, *Psychol. Bull.*, 1944, 41, 345–376. Special attention is called to the style of numbers and abbreviations (p. 353). The term 'experimenter,' with its various forms, singular, plural, and possessive singular and plural, is abbreviated *E*, *Es*, *E*'s, and *Es*'; analogous abbreviations are used for the terms 'subject' (*S*) and 'observer' (*O*). All manuscripts must be typed double-spaced throughout, including tables, footnotes, references, and titles for figures. Tables, figures, footnotes, references, and titles for figures must be placed on separate pages at the end of the manuscript.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879.

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
HUMAN RESOURCES RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *Stanford University*
Donald B. Lindsley, *Northwestern University*
Neal E. Miller, *Yale University*
Edwin B. Newman, *Harvard University*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Margaret K. Harlow, Managing Editor

Volume 41, 1951

PUBLISHED MONTHLY
BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STS., LANCASTER, PA.
AND 1515 MASSACHUSETTS AVE., N. W., WASHINGTON 5, D. C.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879
Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

LANCASTER PRESS, INC., LANCASTER, PA.

CONTENTS OF VOLUME 41

ADAMS, J. A. See LEWIS, D.

ALBERTS, ELIZABETH, AND EHRENFREUND, D. **Transposition in Children as a Function of Age** 30

ALEXANDER, L. T. **The Influence of Figure-Ground Relationships in Binocular Rivalry** 376

AMMONS, R. B. **Effect of Distribution of Practice on Rotary Pursuit "Hits"** 17

AMMONS, R. B. **Effects of Pre-practice Activities on Rotary Pursuit Performance** 187

BAKER, KATHERINE E., WYLIE, RUTH C., AND GAGNÉ, R. M. **The Effects of an Interfering Task on the Learning of a Complex Motor Skill** 1

BARBER, J. L., AND GARNER, W. R. **The Effect of Scale Numbering on Scale-Reading Accuracy and Speed** 298

BARON, M. R. See BROWN, J. S.

BENDIG, A. W. **The Effect of Reinforcement on the Alternation of Guesses** 105

BERSH, P. J. **The Influence of Two Variables upon the Establishment of a Secondary Reinforcer for Operant Responses** 62

BILODEAU, E. A. See BROWN, J. S.

BILODEAU, I. MCD., AND SCHLOSBERG, H. **Similarity in Stimulating Conditions as a Variable in Retroactive Inhibition** 199

BIRCH, H. G., AND RABINOWITZ, H. S. **The Negative Effect of Previous Experience on Productive Thinking** 121

BIRREN, J. E., CASPERSON, R. C., AND BOTWINICK, J. **Pain Measurement by the Radiant Heat Method: Individual Differences in Pain Sensitivity, the Effects of Skin Temperature, and Stimulus Duration** 419

BOTWINICK, J. See BIRREN, J. E.

BRAMMER, L. M. **Sensory Cues in Pitch Judgment** 336

BROWN, J. S., BILODEAU, E. A., AND BARON, M. R. **Bidirectional Gradients in the Strength of a Generalized Voluntary Response to Stimuli on a Visual-Spatial Dimension** 52

BROWN, J. S., KALISH, H. I., AND FARBER, I. E. **Conditioned Fear as Revealed by Magnitude of Startle Response to an Auditory Stimulus** 317

BROWN, W. L. See FOSMIRE, F. R.

BRUCE, R. H., AND LOW, F. N. **The Effect of Practice with Brief-Exposure Techniques upon Central and Peripheral Visual Acuity and a Search for a Brief Test of Peripheral Acuity** 275

BRUNSWIK, E., AND HERMA, H. **Probability Learning of Perceptual Cues in the Establishment of a Weight Illusion** 281

BUNCH, M. E. See MARX, M. H.

CASPERSON, R. C. See BIRREN, J. E.

DENNY, M. R., AND DUNHAM, M. D. **The Effect of Differential Nonreinforcement of the Incorrect Response on the Learning of the Correct Response in the Simple T-Maze** 382

DUNLAP, R. See LEUBA, C.

EHRENFREUND, D. See ALBERTS, ELIZABETH.

ENINGER, M. U. **The Rate of Learning a Tone–No-Tone Discrimination as a Function of the Tone Duration at the Time of the Choicepoint Response** 440

ENINGER, M. U. **The Role of Irrelevant Drive Stimuli in Learning Theory** 446

ESCHENBACH, A. See RETHLINGSHAFER, DOROTHY.

FAGAN, C. A., AND NORTH, A. J. **A Verification of the Guessing-Sequence Hypothesis About Spread of Effect** 349

FARBER, I. E. See BROWN, J. S.

FOSMIRE, F. R. AND BROWN, W. L. **The Effect of Training Procedures on the Relative Strength of Place and Direction Dispositions** 450

GAGNÉ, R. M. See BAKER, KATHERINE E.

GARNER, W. R. **The Accuracy of Counting Repeated Short Tones** 310

GARNER, W. R. See BARBER, J. L.

GRANT, D. A. **Perceptual versus Analytical Responses to the Number Concept of a Weigl-Type Card Sorting Test** 23

GRANT, D. A. See HAKE, H. W.

GUEDRY, F. E. See MANN, C. W.

HAKE, H. W., AND GRANT, D. A. **Resistance to Extinction and the Pattern of Reinforcement: II. Effect of Successive Alternation of Blocks of Reinforced and Unreinforced Trials upon the Conditioned Eyelid Response to Light** 216
HAKE, H. W., GRANT, D. A., AND HORNSETH, J. P. **Resistance to Extinction and the Pattern of Reinforcement: III. The Effect of Trial Patterning in Verbal "Conditioning"** 221
HEISE, G. A. See MILLER, G. A.
HELMICK, J. S. **Pursuit Learning as Affected by Size of Target and Speed of Rotation** . 126
HERMA, H. See BRUNSWIK, E.
HOCHBERG, J. E., TRIEBEL, W., AND SEAMAN, G. **Color Adaptation under Conditions of Homogeneous Visual Stimulation (Ganzfeld)** 153
HORNSETH, J. P. See HAKE, H. W.
HOUSE, BETTY J. See ZEAMAN, D.
HOWES, D. H. AND SOLOMON, R. L. **Visual Duration Threshold as a Function of Word-Probability** 401
HUBBARD, W. R. **Secondary Reinforcement of a Simple Discrimination in Human Beings** 233
HUNT, J. MCV. See WINNICK, WILMA A.
HURVICH, L. M. See JAMESON, DOROTHEA.
IRION, A. L., AND WHAM, DOROTHY S. **Recovery from Retention Loss as a Function of Amount of Pre-recall Warming-up** 242
JAMESON, DOROTHEA AND HURVICH, L. M. **Use of Spectral Hue-Invariant Loci for the Specification of White Stimuli** 455
JARVIK, M. E. **Probability Learning and a Negative Recency Effect in the Serial Anticipation of Alternative Symbols** 291
JOHNSON, E. P., AND RIGGS, L. A. **Electroretinal and Psychophysical Dark Adaptation Curves** 139
KALISH, H. I. See BROWN, J. S.
KENDLER, H. H., AND LEVINE, S. **Studies of the Effect of Change of Drive: I. From Hunger to Thirst in a T-Maze** 429
KIMBLE, G. A. **Behavior Strength as a Function of the Intensity of the Hunger Drive** 341
LEUBA, C., AND DUNLAP, R. **Conditioning Imagery** 352
LEVINE, S. See KENDLER, H. H.
LEWIS, D., MCALLISTER, DOROTHY E., AND ADAMS, J. A. **Facilitation and Interference in Performance on the Modified Mashburn Apparatus: I. The Effects of Varying the Amount of Original Learning** 247
LEWIS, D. See MCALLISTER, DOROTHY E.
LEYZOREK, M. **Two-Point Discrimination in Visual Space as a Function of the Temporal Interval between the Stimuli** 364
LIBERMAN, A. M. **A Comparison of Transfer Effects during Acquisition and Extinction of Two Instrumental Responses** 192
LICHTEN, W. See MILLER, G. A.
LOW, F. N. See BRUCE, R. H.
MANN, C. W., GUEDRY, F. E., AND RAY, J. T. **Post-rotational Perception of Apparent Bodily Rotation** 114
MANN, C. W., AND PASSEY, G. E. **The Perception of the Vertical: V. Adjustment to the Postural Vertical as a Function of the Magnitude of Postural Tilt and Duration of Exposure** 108
MARKS, M. R. **Problem Solving as a Function of the Situation** 74
MARX, M. H., AND BUNCH, M. E. **New Gradients of Error Reinforcement in Multiple-Choice Human Learning** 93
MAUNEY, J. See WORCHEL, P.
MCALLISTER, DOROTHY E., AND LEWIS, D. **Facilitation and Interference in Performance on the Modified Mashburn Apparatus: II. The Effects of Varying the Amount of Interpolated Learning** 356
MCALLISTER, DOROTHY E. See LEWIS, D.
MEYER, D. R. **Food Deprivation and Discrimination Reversal Learning by Monkeys** .. 10
MEYER, D. R. **The Effects of Differential Rewards on Discrimination Reversal Learning by Monkeys** 268
MILLER, G. A., HEISE, G. A., AND LICHTEN, W. **The Intelligibility of Speech as a Function of the Context of the Test Materials** 329

MONTGOMERY, K. C. **An Experimental Investigation of Reactive Inhibition and Conditioned Inhibition** 39
NORTH, A. J. See FAGAN, C. A.
NOTTERMAN, J. M. **A Study of Some Relations among Aperiodic Reinforcement, Discrimination Training, and Secondary Reinforcement** 161
PAINTAL, A. S. **A Comparison of the Galvanic Skin Responses of Normals and Psychotics** 425
PASSEY, G. E. See MANN, C. W.
RABINOWITZ, H. S. See BIRCH, H. G.
RAY, J. T. See MANN, C. W.
RETHLINGSHAFER, DOROTHY, ESCHENBACH, A., AND STONE, J. T. **Combined Drives in Learning** 226
RIGGS, L. A. See JOHNSON, E. P.
ROSENBAUM, G. **Temporal Gradients of Response Strength with Two Levels of Motivation** 261
SALTZMAN, I. J. **Delay of Reward and Human Verbal Learning** 437
SCHLOSBERG, H. See BILODEAU, I. MCD.
SEAMAN, G. See HOCHBERG, J. E.
SOLOMON, R. L. See HOWES, D. H.
STONE, G. R., AND WALTER, N. **The Effect of Negative Incentives in Serial Learning: VI. Response Repetition as a Function of an Isolated Electric Shock Punishment** .. 411
STONE, J. T. See RETHLINGSHAFER, DOROTHY.
TAYLOR, JANET A. **The Relationship of Anxiety to the Conditioned Eyelid Response** .. 81
TEEL, K., AND WEBB, W. B. **Response Evocation on Satiated Trials in the T-Maze** ... 148
TRIEBEL, W. See HOCHBERG, J. E.
WALTER, N. See STONE, G. R.
WEBB, W. B. See TEEL, K.
WINNICK, WILMA A., AND HUNT, J. MCV. **The Effect of an Extra Stimulus upon Strength of Response during Acquisition and Extinction** 205
WHAM, DOROTHY S. See IRION, A. L.
WORCHEL, P., AND MAUNEY, J. **The Effect of Practice on the Perception of Obstacles by the Blind** 170
WYLIE, RUTH C. See BAKER, KATHERINE E.
YAMAGUCHI, HARRY G. **Superthreshold Reaction Potential ($s\bar{E}R$) as a Function of Experimental Extinction (n)** 391
ZEAMAN, D., AND HOUSE, BETTY J. **The Growth and Decay of Reactive Inhibition as Measured by Alternation Behavior** 177

Journal of
Experimental Psychology

VOL. 41 NO. 6 JUNE 1951

SUPERTHRESHOLD REACTION POTENTIAL (${}_{S}\bar{E}_{R}$) AS A FUNCTION OF EXPERIMENTAL EXTINCTION (n) [1]

BY HARRY G. YAMAGUCHI

Department of Psychology, Institute of Human Relations, Yale University

A number of studies have been made on experimental extinction as a function of various aspects of training, such as the number of reinforcements employed (**5**, **6**). However, no study has determined

$$ {}_{S}\bar{E}_{R} = f(n) . $$

This relationship is presumably a functional behavioral law and is important as such. Moreover, at the present time it has the additional importance that it is required to mediate the determination of drive (D) as a function of the number of hours of food privation, i.e.,

$$ D = f(h) . $$

METHOD

Subjects.—The *S*s were 199 albino rats purchased from the Albino Farms, Redbank, New Jersey. At the beginning of training they ranged from 7 to 8 months of age with a mean food satiation weight of 247 gm. The animals were housed in group cages 16 in. square and 10 in. high, 10 *S*s per cage, but were fed by pairs in smaller cages. The diet consisted of Purina Laboratory Chow supplemented by pellets of Gaines Krunchons saturated with cod-liver oil.

Apparatus.—The apparatus was similar to the modified Skinner box used in the study by Felsinger *et al.* (1). This box was shielded against external sounds and was ventilated by suction. In the interior a plywood partition separated the anterior half, containing an animal restriction chamber with a 7.5-w. bulb, from the posterior half of the box containing the operating mechanisms. The partition was sheathed with sheet-metal on the side facing the animal chamber. A 7.5-in. observation porthole was located above the anterior half. The animal restriction chamber was 8.5 in. long, 6 in. wide, and 5 in. high, with sides of wood and a clear glass top. This chamber contained a vertically sliding shutter of sheet-metal located .5 in. from the wall of the apparatus chamber. The shutter was lowered by a cord which passed through a small hole in the wall of the box, and was raised by an external spring mechanism.

The mechanisms in the posterior chamber of the Skinner box consisted of the manipulandum and pellet-dispensing units. The manipulandum was made of $\frac{1}{4}$-in. brass tubing and projected $\frac{1}{2}$ in. horizontally into the animal chamber through a slit in the partition. It was pivoted to allow lateral movement to the left against spring tension. When a leftward pressure of approximately 7.5 gm. was applied with a movement of $\frac{1}{16}$ in., a circuit containing two electromagnets was closed through a contact point on the posterior end of the manipulandum

[1] The present study is a portion of a thesis presented to the Faculty of the Graduate School of Yale University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The investigation also represents a part of the coordinated research program of the Institute of Human Relations. The writer is deeply indebted to Professors Clark L. Hull and Mark A. May for their generous advice and criticism.

tube. One of these magnets released a locking device on the manipulandum carriage and permitted the manipulandum to retract behind the partition. The other magnet simultaneously released a plunger activated by a weight and pulley system to force a pellet of food partially through the stationary pellet-holding tube. This tube extended to the slit in the panel and was mounted $\frac{1}{16}$ in. to the right of the manipulandum. It was formed of sheet spring brass and was constructed with a tapered bore. The forward movement of the pellet plunger was arrested $\frac{1}{16}$ in. short of ejecting the pellet into the animal chamber, and the constrictive tension of the pellet tube held the pellet until it was seized by the animal. The pellet-dispensing mechanism was gravity loaded through a glass tube which descended into the box through a hole in the top. Friction and the tapered bore of the pellet tube stopped the forward descent of the pellet dropped into the pellet-dispensing mechanism. The pellet was not visible to *S* until the pellet plunger had been activated.

Pellets were made of a dry mix of two parts of finely ground Laboratory Chow to one part flour. Water was added to form a doughy mash which was extruded through a die. The extruded material was rolled into smooth sticks and was thoroughly dried. Quarter-inch sections were cut after the dry sticks had first been made pliable by being placed in a box lined with wet paper towels. These humidified pellets were approximately $\frac{1}{4}$ in. in diameter and averaged .2 gm. with a 50 per cent water content by weight.

Reaction latencies were measured in .01 sec. by a Telechron timer and were estimated to .001 sec. to reduce errors in the scaling of latency data (4, p. 244). The timer started when the shutter-raising spring was tripped, and stopped when the electric contact produced through a manipulandum movement was established. To minimize variability of reaction latency, the shutter was lifted only if *S* was in the posterior half of the animal chamber and was facing the shutter. Sighting points and lines were painted on the observation porthole to minimize subjectivity in observing this rule of procedure.

Procedure.—One of the experimental variables necessarily studied in this investigation was the number of reinforcements given before experimental extinction of the manipulandum response. All learning and extinction trials were performed 24 hr. after satiation. The number of reinforcements given before extinction were 2, 4, 9, 21, and 88. Five different groups of *Ss* were run corresponding to the five different numbers of reinforcements. The number of *Ss* in each group is presented in Table I.

The *Ss* were randomly assigned to the various groups. Assignment to living cages was also on a randomized basis. Group Vb was run prior to the other groups and consisted of 51 *Ss* identical in age and source with those of the other groups. Groups Va and Vb were given identical experimental treatment with one exception: in addition to 88 reinforcements at 24 hr. of hunger, each *S* in Group Vb was given one reinforced trial after each of the three different food deprivation intervals of 3, 48, and 72 hr. These test trials occurred in counterbalanced order and were distributed between Trials 72–73, 76–77, and 80–81 to minimize any interference effects of one test trial on another. Eight training trials intervened between the last of these trials and extinction.

TABLE I

NUMBER OF *Ss* IN GROUPS WHICH VARIED IN THE NUMBER OF REINFORCEMENTS GIVEN BEFORE EXTINCTION

Group	No. of Reinforcements (*N*)	No. of *Ss*	Original No. of *Ss*
I	2	34	40
II	4	31	40
III	9	34	40
IV	21	34	40
Va	88	28	35
Vb	88	38	51

The *Ss* were not kept on a maintenance cycle of daily feedings, but were maintained on a 5-day experimenting cycle of 72 hr. of hunger followed by 48 hr. of feeding ad libitum. This maintenance cycle was used in order to effect an identity in maintenance procedure between this experiment and a related study (7) in which the hours of hunger ranged up to 72.

The extreme degree of hunger to which the *Ss* were repeatedly exposed was mainly responsible for the reduction of the original size of the groups, as indicated in Table I. Of an original total of 246 *Ss*, 47, or 19 per cent, were lost either through starvation or incapacitating injuries received in the group living cages.

Because of the 5-day experimenting cycle, the *Ss* were given only one training session every five days. Each group was divided into five subgroups so that one-fifth of each group was run each day.

Habituation to maintenance cycle.—The 5-day maintenance cycle was established by systematically increasing the starvation interval by 24 hr. over a period of 12 days. Enforced satiation was initiated at the beginning of habituation to the maintenance cycle. To satiate the animals, an ample supply of dry 1-

to 3-in. lengths of pellet material was placed in the feeding cages approximately 1 hr. before the feeding period ended. An animal was considered satiated when it refused to accept additional sticks of food offered by hand, or when it did not consume food within 3 to 5 min. after taking it.

Habituation to apparatus.—Habituation to the apparatus and the pellet-dispensing mechanism was conducted under 24-hr. hunger and was carried out simultaneously with the habituation to the maintenance cycle. On the first day a 10-min. period was allowed each *S* for exploratory activity in the animal restriction chamber. On the second day, Laboratory Chow was placed on the floor of the chamber, and the animals were allowed to eat and explore for 10 min. During the next two days, ten Skinner-box pellets were placed on the floor at the posterior end of the animal chamber. The technique used in the last stage of habituation was similar to that used by Perin (5) and consisted in presenting pellets through the pellet-holding tube, the plunger mechanism being operated by *E*. The beginning of this phase coincided with the end of the habituation to the maintenance cycle. Therefore, one habituation session was given every 5 days. Four pellets were presented on each of seven sessions until a total of 28 pellets had been given.

Training the manipulandum response.—Five days after the last habituation trial the manipulandum was introduced for the first time, and training on the bar-movement response commenced. Ten seconds after the apparatus had been cocked and baited, the shutter was raised. This was done, however, only if *S* was facing the shutter and some portion of its body was in the posterior half of the animal compartment. The raising of the shutter was delayed until these conditions were met, as judged by *E* with the aid of the sighting lines on the observation porthole. When moved leftward under sufficient pressure, the manipulandum retracted, the timer stopped, and a pellet was presented through the delivery tube. The shutter was then lowered and the apparatus cocked and baited for the next trial. One session of four massed trials was given every 5 days until the required number of reinforcements was attained. In the case of Groups I, III, and IV the required numbers were not multiples of four, so that the last training session consisted of less than four trials for these groups.

To prevent temporal conditioning, the trials given within each training session were separated by varied intervals. The beginning of a trial was 5, 10, or 15 sec. after the animal had consumed the pellet of the previous trial. For the first trial of each session, the cocking and baiting of the apparatus marked the beginning of the trial-separation interval. Beginning with the first training trial, the sequence of these intervals was 10, 5, and 15 sec., the sequence being repeated for the subsequent trials.

Extinction.—All *S*s were extinguished by massed trials at 24 hr. of hunger, 5 days after the last training trial. An arbitrary extinction criterion of 2 min. was used. The trials were separated by the same intervals and sequence of intervals as the training trials. Since no reinforcement was presented, however, the trial-separation intervals were measured from the cocking of the apparatus.

Results

Homogeneity of Groups

Although the *S*s were randomly assigned to the various groups, it can be seen from Table I that the proportion of nonsurvivors varied somewhat between groups. The differences, however, were not significant since the chi-square is only 2.97 for 5 *df*; a chi-square as large or larger could occur 70 per cent of the time by chance.

To establish the homogeneity of the groups with respect to learning ability, an *F* test was made of the manipulandum latency data at Trial 2, the last trial all groups had in common. All groups were at 24-hr. hunger. With the latencies transformed into log units to eliminate the positive skew typically found in such data, a two-component analysis of variance resulted in a between-groups variance of .0576 and a within-groups variance of

TABLE II

Latency of Response (st_R) in Seconds for All Groups after One Reinforcement

	Reinforcement Group					
	I	II	III	IV	Va	Vb
Number of reinforcements	2	4	9	21	88	88
Median latency	8.75	8.00	7.36	10.07	9.46	5.02
Mean log latency	0.950	0.925	0.963	0.952	1.001	0.877

.1303. Since the variance between groups is less than the error term, the lack of significant differences between the groups is evident. Although the within-groups variance is larger than the between-groups variance, the ratio of these two variances was not significantly different from 1. Table II presents the median latencies and mean log latencies at Trial 2.

Number of Reactions to Extinction (n) *as a Function of the Number of Reinforcements* (N)

The mean and median numbers of extinction responses of the various reinforcements are presented in Table III. Table IV gives the critical ratios and the significances of the differences between medians. The data of the two 88-reinforcement groups (Va and Vb) have been pooled, since the difference in the median number of reactions to extinction between the two subgroups is not greater than could be expected by chance. With a median n of 23 for Group Va and of 20 for Group Vb, the t of the difference is only .739 with 64 *df*.

The form of the relationship of n to N is shown in Fig. 1, where the median n values of Table III have been plotted as a function of N. The curve of best fit by the criterion of least squares is based on the equation

$$n = 23.00(1 - 1.306\,N^{-.749}) - .11. \quad (1)$$

In a somewhat similar situation, Williams (6) reports the following fitted equation:

$$n = 62 - 62 \times 10^{-.017\,N}. \quad (2)$$

Perin (5) obtained a better fit to the Williams data by using the equation

$$n = 66(1 - 10^{-.018\,N}) - 4. \quad (3)$$

A comparison of equation (1) with the equations of Williams and Perin reveals that equation (1) predicts an

TABLE III

EFFECT OF NUMBER OF REINFORCEMENTS DURING TRAINING ON NUMBER OF EXTINCTION RESPONSES

Group	No. of Reinforcements	*Ss*	No. of Extinction Responses		
			Mean	Median	$\sigma_{Mdn.}$
I	2	34	5.62	5.00	.85
II	4	31	17.42	12.00	1.83
III	9	34	24.91	18.00	3.49
IV	21	34	29.56	19.00	5.80
Va+Vb	88	66	24.45	22.00	2.17

TABLE IV

SIGNIFICANCE OF DIFFERENCES IN MEDIAN NUMBER OF REACTIONS TO EXTINCTION BETWEEN PAIRED GROUPS

Paired Groups	Difference	Critical Ratio	*P*
2 reinf. vs. 4	7	3.468	<.01
2 reinf. vs. 9	13	3.614	<.01
2 reinf. vs. 21	14	2.390	.02
2 reinf. vs. 88	17	7.301	<.01
4 reinf. vs. 9	6	1.521	.13
4 reinf. vs. 21	7	1.152	.25
4 reinf. vs. 88	10	3.525	<.01
9 reinf. vs. 21	1	0.148	.88
9 reinf. vs. 88	4	0.973	.33
21 reinf. vs. 88	3	0.485	.63

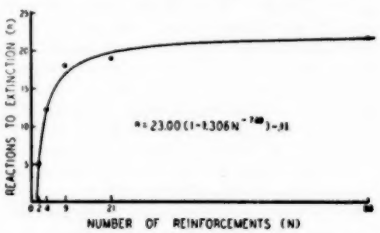

FIG. 1. Graphic representation of the functional relationship of median number of reactions for extinction to the number of reinforcements (N). The fitted curve was plotted from the equation shown.

asymptote of 22.93 extinction reactions, whereas equations (2) and (3) predict an asymptote of 62; and equation (1) is an inverse power function which approaches its asymptote relatively more rapidly than do the other two equations. The first difference may be essentially due to the use of a 2-min. extinction criterion instead of the 5-min. criterion used in the Williams experiment, and to the possibility that a greater amount of work was involved in the lateral manipulandum movement response than in the bar depression response in the Williams Skinner-box. The second difference may presumably be attributed to differences in reinforcement (3) caused by (*a*) the use of humidified pellets of approximately twice the dry weight of the hard pellets used by Williams, and (*b*) subjecting the animals of the present study to 3 days of food privation in their maintenance schedule in contrast to the 1 day of hunger involved in the conventional daily feeding cycle in the Williams experiment.

Reaction Potential ($_sE_R$) *as a Function of the Number of Reinforcements* (N)

To determine the relationship between reaction potential and number of reinforcements, we employed the Thurstone scaling technique modification elaborated by Hull *et al.* (4), using the pooled learning data of the 88-reinforcement group on a selected number of trials. A selection of trials was necessary to reduce the labor of computation, and was done in such a manner that the selected N points were relatively close together at the beginning of training where $_sE_R$ would increase rapidly, but were more widely separated as the number of reinforcements increased. No trials between $N = 72$ and $N = 81$, where the animals of Group Vb were tested at different drive levels, were included in the scaling. The 25 trials thus selected for scaling are listed in the first column of Table V.

Since the animals were trained in sessions of four consecutive trials with successive sessions separated by 5 days, it was necessary to determine whether the grouping of trials and the spacing of these groups had produced any consistent distortions in the learning latencies which would affect the scaling results. Since a lawful relationship between reaction potential and median latency of response had been found by Gladstone *et al.* (2), the median latencies were examined. By fitting a least-squares curve to the plot representing the relationship between latency of response ($_st_R$) and number of reinforcements (N), the following equation was obtained:

$$_st_R = 7.12\,(N + .157)^{-.645} + .503\,. \quad (4)$$

The mean square deviation of this curve is .0240 sec.^{2}. The fit was such that 44 of the 88 latency values were above the curve, and 44 were below. A chi-square test of homogeneity was made to determine whether there were any significant departures from the equal over-all proportionality of above versus below values obtained for first, second, third, and fourth within-sessions trials. A chi-square of 1.091 with 3 *df* was obtained; a chi-square equal to or greater than this could happen 82 times in 100 by chance, which indicates that the ordinal position of a trial within a four-trial group did not produce significant distortions from the general tendency of median latency values to decrease with an increase in the number of reinforcements.

Since the use of the normal probability table in the scaling of latency

data produces distortions (8), the table appropriate for the leptokurtic nature of the latency dispersions was used. To test for the appropriateness of the use of this table, a comparison of the direct versus indirect scale separations was made. Since the distortion was only 1.9 per cent with a nonsignificant t of .23 for the difference between direct and indirect determinations, it is apparent that the use of the leptokurtic table was effective.

The ${}_sE_R$ values obtained by the scaling procedure are presented in Table V, with the relationship between ${}_sE_R$ and N presented in Fig. 2. The curve of best fit is represented by the equation

$$ {}_sE_R = 7.58[1 - (N + 1)^{-.201}] + .045 . \quad (5) $$

TABLE V

ORDINAL NUMBER (N) OF SELECTED REINFORCEMENTS, THE PARALLEL REACTION POTENTIALS (${}_sE_R$), AND THE CORRESPONDING MEDIAN REACTION LATENCIES (${}_st_R$)

Ordinal No. (N) of Selected Reinforcements	Reaction Potential (${}_sE_R$)	Median Reaction Latency (${}_st_R$)
0	0.000	24.069
1	1.126	7.086
2	1.825	4.893
3	1.792	4.374
4	1.846	3.992
6	2.625	2.620
8	2.571	2.296
10	3.263	1.578
12	3.067	1.919
14	3.467	1.428
16	3.253	1.749
18	3.491	1.410
20	3.417	1.481
24	3.487	1.226
28	3.620	1.226
32	3.848	1.242
36	3.815	1.330
40	4.009	1.255
44	3.944	1.136
48	4.186	1.050
52	4.202	1.032
60	4.385	0.972
68	4.559	0.879
84	4.696	0.984
87	4.422	0.934

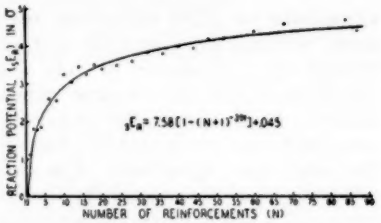

FIG. 2. Graphic representation of the functional relationship between reaction potential (${}_sE_R$) and the number of reinforcements (N) based on the 66 Ss of the 88-reinforcement group. The fitted curve was plotted from the equation shown.

The mean square deviation of this curve is $.0234\sigma^2$. This is the type of equation rejected for this function in a somewhat similar situation by Gladstone *et al.* (2, pp. 521–522).

The relationship between the response latency and reaction potential values of Table V is shown in Fig. 3. The equation of best fit is

$$ {}_sE_R = 6.26\ {}_st_R^{-.353} - 1.969 , \quad (6) $$

with a mean square deviation of $.0209\sigma^2$. This is the same type of equation as that obtained by Gladstone *et al.* (2), but the constants are all appreciably different.

Because the ${}_sE_R$ values in the second column of Table V were based on an arbitrary zero at $N = 0$, the determination of the threshold and of absolute zero was highly desirable to provide a more adequate basal point of reference. In the study by Gladstone *et al.* (2, pp. 513 ff.), a tentative absolute zero (Z) and reaction threshold (${}_sL_R$) were determined on the basis of the number of subthreshold reinforcements required to raise the response strength above threshold.

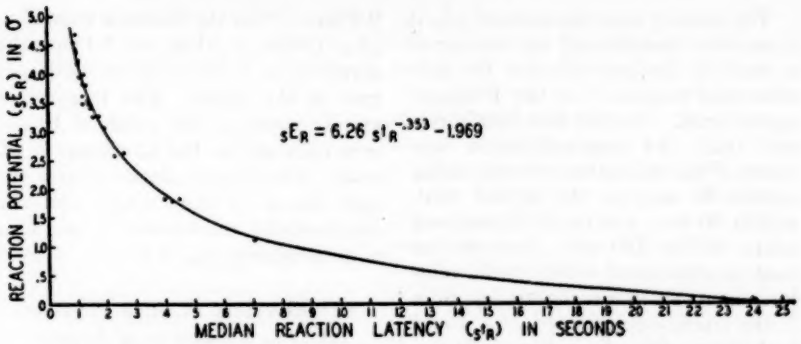

FIG. 3. Graphic representation of the relationship of reaction potential ($_sE_R$) as a function of median reaction latency ($_{st}R$) based on the learning data of the 66 Ss of the 88-reinforcement group. The fitted curve was plotted from the equation shown.

It was found that both the mean number and the median number of subthreshold reinforcements were slightly less than 2. The integral value of 2 was used, and the curve representing the relationship between $_sE_R$ and N was extrapolated (2, p. 512) to determine the distance in $_sE_R$ between Z and the first superthreshold trial. By the equation

$$_sE_R = 3.084(1 - 10^{-.0305\,N}) + .04, \quad (7)$$

Z was found to be $-.426\sigma$. The threshold was defined as .5 less than the number of subthreshold reinforcements, and was found to be $-.071\sigma$, or $.355\sigma$ above Z.

Since the habituation procedure used in the present experiment did not involve the manipulandum as in the above-mentioned study, it is necessary to determine Z and $_sL_R$ by another procedure. We propose, then, that Z and $_sL_R$ be defined in terms of latency values: (*a*) Z is the $_sE_R$ value associated with infinite $_{st}R$; and (*b*) $_sL_R$ is the $_sE_R$ value associated with whatever arbitrary latency criterion is used in defining the threshold in instrumental conditioning, e.g., the 5-min. extinction criterion of the Williams and Perin experiments, the latency interval before baiting the manipulandum in the Felsinger experiment (1), and so forth. By the use of an appropriate equation the latency definitions of Z and $_sL_R$ can be transformed into $_sE_R$ units.

The equivalence of the latency determination procedure and the subthreshold reinforcement method can be demonstrated in the case of the Gladstone data. The equation which enables the transforming of $_{st}R$ into $_sE_R$ is (2):

$$_sE_R = 2.845\ (_{st}R)^{-.433} - .599. \quad (8)$$

Substituting ∞ in this equation yields a value of $-.599\sigma$ for Z. The difference between this value of Z and that of $-.426\sigma$ yielded by the subthreshold reinforcement procedure is not significant.[2]

[2] For a rough estimate of the significance of the difference between the two Z values, the square root of the sum of the mean square deviations of the equations employed was taken as the standard error of the difference. The mean square deviation of equation (7) was $.0275\sigma^2$, and that of equation (6) was $.0143\sigma^2$. The square root of the sum being .204, the difference of .173 between the Z values produces a nonsignificant critical ratio of .848.

The latency determination of ${}_sL_R$ is somewhat complicated by the use of a variable latency criterion for subthreshold responses in the Felsinger experiment. On the first subthreshold trial the manipulandum was baited if the animal had not responded within 30 sec.; on the second trial, within 60 sec.; and on all subsequent trials, within 180 sec. Because the median number of subthreshold reinforcements given was slightly less than 2, the latency-determined ${}_sL_R$ would lie between the ${}_sE_R$ value of 30 sec. and that of 60 sec. By equation (8), ${}_sL_R$ should fall somewhere between $.049\sigma$ (30 sec.) and $-.205\sigma$ (60 sec.). Since the value of ${}_sL_R$ determined by the subthreshold reinforcement procedure was $-.071\sigma$, a fair consistency of the results of two different procedures is again demonstrated.

For the present experiment, by using equation (6) and defining Z as ∞ ${}_{st}R$ and ${}_sL_R$ as 120 sec., the arbitrary criterion of extinction, we obtained the following results:

$$Z = -1.969\sigma\,,$$

$${}_sL_R = -.814\sigma\,.$$

Since .814 must be added to all the values in the second column of Table V to produce superthreshold reaction potential (${}_s\dot{E}_R$), the equation for ${}_s\dot{E}_R$ can be based on equation (5) and written as

$${}_s\dot{E}_R = 7.58[1 - (N+1)^{-.201}] + .859\,. \quad (9)$$

Similarly, the equation for absolute reaction potential (${}_s\dot{E}_R$) is

$${}_s\dot{E}_R = 7.58[1 - (N+1)^{-.201}] + 2.014\,. \quad (10)$$

The range of reaction potential from Z to the maximum is given by equation (10) as $7.58\sigma + 2.014\sigma$, or 9.594σ. With the distance from Z to ${}_sL_R$ $1.969\sigma - .814\sigma$, or 1.155σ, the threshold is 1.155/9.594 or 12.05 per cent of the range. This is approximately equal to the value of 10 per cent obtained in the Gladstone *et al.* study, which again demonstrates the equivalence of the latency and the subthreshold reinforcement methods of determining ${}_sL_R$ and Z.

Superthreshold Reaction Potential (${}_s\dot{E}_R$) *as a Function of Experimental Extinction* (n)

The values used in the equation-fit for the determination of the relationship,

$${}_s\dot{E}_R = f(n)\,,$$

are shown in Table VI. In this connection it is to be noted that the n values were empirically obtained for each group, but the ${}_s\dot{E}_R$ values were calculated by substituting the appropriate N values in equation (5), which was derived by scaling the learning data of the 88-reinforcement group. A graphic representation of the values involved is shown by the circles in

TABLE VI

MEDIAN NUMBER OF REACTIONS TO EXTINCTION (n) WITH THE CORRESPONDING SUPERTHRESHOLD REACTION POTENTIAL (${}_s\dot{E}_R$) VALUES FOR VARIOUS NUMBERS OF REINFORCEMENTS

The superthreshold reaction potential values were calculated by the equation,

$${}_s\dot{E}_R = 7.58[1 - (N+1)^{-.201}] + .859\,.$$

No. of Reinforcements (N)	No. of Reactions to Extinction (n)	Superthreshold Reaction Potential (${}_s\dot{E}_R$)
2	5	2.361
4	12	2.954
9	18	3.667
21	19	4.367
88	22	5.364

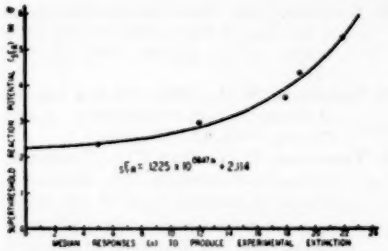

FIG. 4. Graphic representation of the relationship between superthreshold reaction potential ($_sE_R$) and median number of responses (n) to produce experimental extinction. The $_sE_R$ values were calculated by substituting the the proper values of number of reinforcements (N) in equation 9. The curve was plotted from the equation (11) shown.

Fig. 4. The equation fit to these data is,

$$_sE_R = .1225 \times 10^{.0647\,n} + 2.114\,, \quad (11)$$

and is represented by the smooth curve drawn among the circles. The mean square deviation of this fit is $.020\sigma^2$. The equation, while not a very close fit, evidently indicates a first approximation to the law we have been seeking.

The curve shown in Fig. 4 intersects the ordinate at $_sE_R = 2.236$, not at zero as might be expected on the basis of the previous discussion concerning $_sE_R$, $_sL_R$, and the criterion of extinction. If the assumption that $n = 0$ when $_sE_R = 0$ is valid, experimental study of the region of n below the range obtained in the present study may reveal an inflection in the lower region of the function shown in Fig. 4. Equation (11), therefore, is of questionable accuracy below $n = 5$ or $_sE_R = 2.361$.

SUMMARY

Five groups of albino rats totaling 199 *S*s were trained and extinguished on a manipulandum movement response in a modified Skinner box. Both learning and extinction were carried out 24 hr. after food satiation, and the several groups were extinguished after 2, 4, 9, 21, and 88 quasi-distributed reinforcements. The relationship of number of extinction responses to reaction potential was determined by the following steps:

1. The number of reactions to extinction (n) with a 120-sec. criterion was found to be positively related to the number of reinforcements (N) (Eq. [1]).

2. The relationship of reaction potential ($_sE_R$) to number of reinforcements was found by applying a paired-comparison scaling technique to the learning data of the 66 *S*s of the 88-reinforcement group (Eq. [5]).

3. The relationship of reaction potential to the median reaction latency ($_st_R$) values of the trials included in the $_sE_R$ scaling yielded a negative concave-upward curve (Eq. [6]).

4. The equivalence of latency-determined $_sE_R$ values of absolute zero (Z) and reaction threshold ($_sL_R$) to those determined by subthreshold reinforcements was found.

5. The relationship of superthreshold reaction potential ($_sE_R$) to the number of extinction responses (n) is represented by a positive concave-upward curve (Eq. [11]).

(Manuscript received June 3, 1950)

REFERENCES

1. FELSINGER, J. M., GLADSTONE, A. I., YAMAGUCHI, H. G., AND HULL, C. L. Reaction latency ($_st_R$) as a function of the number of reinforcements (N). *J. exp. Psychol.*, 1947, **37**, 214–228.

2. GLADSTONE, A. I., YAMAGUCHI, H. G., HULL, C. L., AND FELSINGER, J. M. Some functional relationships of reaction potential ($_sE_R$) and related phenomena. *J. exp. Psychol.*, 1947, **37**, 510–526.

3. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.

4. Hull, C. L., Felsinger, J. M., Gladstone, A. I., and Yamaguchi, H. G. A proposed quantification of habit strength. *Psychol. Rev.*, 1947, **54**, 237–254.

5. Perin, C. T. Behavior potentiality as a joint function of the amount of training and the degree of hunger at the time of extinction. *J. exp. Psychol.*, 1942, **30**, 93–113.

6. Williams, S. B. Resistance to extinction as a function of the number of reinforcements. *J. exp. Psychol.*, 1938, **23**, 506–521.

7. Yamaguchi, H. G. Drive (D) as a function of number of hours of hunger (h). *J. exp. Psychol.*, 1951, **42**. (In press.)

8. Yamaguchi, H. G., Hull, C. L., Felsinger, J. M., and Gladstone, A. I. Characteristics of dispersions based on the pooled momentary reaction potentials ($s\hat{E}_R$) of a group. *Psychol. Rev.*, 1948, **55**, 216–238.

VISUAL DURATION THRESHOLD AS A FUNCTION OF WORD-PROBABILITY [1]

BY DAVIS H. HOWES [2] AND RICHARD L. SOLOMON

Harvard University

Conventional threshold experiments provide a sample of language behavior under highly controlled conditions. However exclusively *E* may be interested in *S*'s vision or hearing, his observations typically are in units of communicative behavior. Of particular interest is the threshold experiment in which a word is exposed tachistoscopically for very brief durations which are gradually lengthened on successive exposures until the word is correctly reported. In such an experiment, the observations are in fundamental units (time) and can be expressed as functions of any measurable properties of behavioral units such as words. Any general property of words, then, should be operative in the word-threshold experiment, and the dependent variable (duration threshold) some function of it.

This paper explores experimentally the function relating duration threshold to the relative frequency with which a word appears in the English language. Word-probability is an inference from relative frequency of occurrence; the only operation necessary to estimate the probability of occurrence of a specific word under any given conditions is to count the frequency of occurrence of such a word relative to the total number of words occurring in a specific sample of language recorded under those conditions. In the present experiment, word-probability will be defined by relative frequency of occurrence of a given word in certain standard word counts for the English language.

Any function which relates visual duration thresholds to word-probability tends to emphasize the *importance of operational analysis in perceptual experiments*. Those perceptual studies measuring the optical conditions under which specific linguistic responses occur are clearly based on *linguistic as well as optical operations*. As the following experiments will show, the probability of a word has an enormous effect on its visual duration threshold; yet the conventional concepts of perception offer no clear interpretation of such a variable. In this paper we shall confine ourselves to a description of experimental data. Their theoretical implications will be discussed in later papers.

Our experimental problem may be phrased as follows: What is the relationship between word-probability and speed of recognition? More specifically, What is the relationship between the relative frequency of occurrence of specific words in the English language and the visual duration threshold necessary for correct report of those words when they are tachistoscopically exposed?

METHOD

Word-frequency.—Word-frequency counts are made by selecting a sample of language behavior (usually written) that contains a given number of words, and then tabulating the number of times that each particular word occurs. The results of several tabulations on large samples of

[1] Drs. E. G. Boring and Leo Postman have graciously read this manuscript and have suggested numerous improvements. This research was supported by the Laboratory of Social Relations, Harvard University.

[2] Now at Tulane University.

written language are incorporated in the Thorndike-Lorge *Teacher's Word Book of 30,000 Words* (4). To minimize the effect of any peculiarities of any individual word counts, the relative frequency of words in the following experiments was estimated in three cases: (1) by the Lorge Magazine Count, (2) by the Lorge-Thorndike Semantic Count, and (3) by the geometric mean of the frequencies in both counts.

The Magazine Count (4) is based on approximately four and one-half million words taken from issues of popular magazines that appeared between 1927 and 1938. The Semantic Count (1) is based on the *Encyclopedia Britannica, Bartlett's Familiar Quotations*, the *Literary Digest*, and books of fiction and nonfiction.[3] Word-frequencies based on the various counts will hereafter be designated as follows: Magazine Count, f_{mg}; Semantic Count, f_{sm}; average of the two counts, f_{av}.

Visual duration threshold.—The 20 *S*s in our experiments were given the following instructions:

> This is an experiment to see how keen your vision is when a word is briefly flashed before your eyes. You will be given words to identify. A number of increasingly long flashes will be presented for each word. I will notify you before I change the word that is being flashed.
>
> The first flashes will be very brief, and you probably will be unable to recognize the word. I will then gradually increase the duration of the flash for the successive trials of that word. After each flash, please state clearly and distinctly whatever you think you saw. There is no objection to guessing; but, if you have no idea what it is, say so.
>
> Before I present each flash, I will say, "Ready." You will then look into the eye-piece and focus on the area between the two orange lines. After the ready signal, do not say anything until you have made your response to the flash. After you have made your response, sit back and relax until I give the ready signal for the next presentation. We will have a recess after every few words, or whenever you say you are tired.
>
> The first few words will be for practice, so that you will become familiar with this procedure.
>
> Do you have any question?

It is important to notice that *S* was not instructed to respond as quickly as possible to the exposure, but to respond whenever he felt ready to respond. The time that *E* measured and controlled was the *duration of the exposure*, not the latency of the *S*'s response.

The words were exposed in a modified Dodge tachistoscope designed and built by Mr. Ralph Gerbrands. The duration of each exposure was controlled to the nearest 10 msec. The range of exposure durations was from 10 to 1,000 msec. The illumination for exposed words was obtained by two preheated fluorescent lamps which could be turned on and off almost instantaneously. The pre-exposure fixation pattern consisted of two horizontal parallel lines of light furnished by neon glow lamps. The pre-exposure illumination was terminated when the word-exposure illumination went on. The *S*, therefore, saw the fixation lines disappear at the onset of the exposure, and a pattern of letters appeared in place of the lines. The flash illumination at the eyepiece was about .07 ft.-candles (Macbeth illuminometer reading), which was sufficiently low that no *S* responded correctly to a word which was exposed for less than 30 msec.

The list of experimental words was typed on an adding-machine roll, which was attached to some rollers at the back of the tachistoscope, and each of the words could be centered, one by one, in the exposure slot in a standard order. The letters were capitals of block type, ¼ in. high, and were exposed approximately 23 in. from *S*'s eyes.

Practice has an enormous effect of this species of tachistoscopy. Four familiar words, all

[3] In these counts the listing of variants of a single root presents a serious problem. Thorndike and Lorge described their system thus: "Regular plurals, comparatives and superlatives, verbs, forms in *s*, *d*, *ed*, and *ing*, past participles formed by adding *n*, adverbs in *ly* that occur less than once in a million words, and rare adjectives formed by adding *n* to names of places are ordinarily counted in under the main word. So are words which have special meanings when capitalized and words usually capitalized which have special meanings when not capitalized. . . . Some adjectives in *ing* or *ed*, and some nouns in *ing*, are entered separately . . ., but, in general, participial adjectives and verbal substantives are included under the word" (4, p. ix). Frequencies of rare adjectives are next to impossible to estimate accurately under this system, and much confusion results from the lack of any apparent reason for the separate listing of any *ing* and *ed* adjectives and the absorption of others to the root. Our own policy was to consider a plural no different from its singular form and to assess the frequency of words ending in *ing* at half the frequency of its root. Only one case arose in which the latter adjustment was necessary.

pertaining to the academic environment (*HARVARD, UNIVERSITY, TEACHER,* and *PROFESSOR*), served as preliminary words to absorb the most striking practice effects, but only a much longer prefatory list could have stabilized the thresholds. To distribute the remaining effects of practice, the standard order in which the words were arranged was counterbalanced for order and staggered for initial position in the list. For example, in a 60-word list used in Experiment I, the word order for successive *S*s followed these patterns (numbers represent words and their standard order): 1, 2, 3 . . . 59, 60; 60, 59, 58 . . . 2, 1; 5, 6, 7 . . . 3, 4; 56, 55, 54 . . . 58, 57; 9, 10, 11 . . . 7, 8; 52, 51, 50 . . . 54, 53; etc. About 20 or 30 thresholds could be measured in an hour. Despite recesses and short rest periods, the last portion of long sessions found the *S*s fatigued, but as thresholds showed a regular decline with practice down to the last word, it would seem that whatever effect fatigue had on thresholds was masked by the large practice effect.

A rhythm was forced by the time required for *E* to record the response to the previous exposure, to set the duration of the next, and to give the ready signal. During this period between flashes (lasting from 15 to 45 sec.) *S* looked around in an experimental room kept at its normal illumination (5 ft.-candles). In this way, *S* was light-adapted to a fairly constant level after each exposure. A brief period (.5 to 1.0 sec.) between the moment *S*'s eyes were sealed against the eyepiece of the tachistoscope and the instant of the flash did permit some dark adaptation to occur, but its degree was unrelated to any properties of the stimulus-words used in our experiments. Permitting *S* to look wherever he wished between exposures also improved his temper during the long session.

Correct responses began to occur with exposures ranging from 30 to 1,000 msec., depending on the individual *S* and on the word being exposed. No rigid schedule of progression from short to long exposures could cope with such an extensive range. For each *S* the exposure time for the first flash was chosen so that it was at least 40 msec. below the very minimum duration at which correct responses could be expected for the particular *S*, word, and stage of practice. Below 200 msec. the step interval used was 10 or 20 msec.; from 200 to 450 msec. the interval was 20 or 30 msec.; and above 450 msec. the step interval was 50 msec. Two exposures were usually given at each flash duration, before the exposure time was lengthened. The number of exposures preceding the threshold criterion ranged from zero to 48. If this variable had any practice effect on thresholds (and none were observed), it would have been to obscure differences between high- and low-threshold words, rather than to exaggerate them; for the former were, on the average, preceded by many more exposures.

Following the ascending method of limits, the flash duration was systematically lengthened until the exposed word was accurately reported. Three successive correct responses was taken for the criterion for a duration threshold. Quantitatively, threshold was defined as the duration halfway between the duration of the longest exposure to which an incorrect response was made and the duration of the three criterion exposures.[4]

A few *S*s' thresholds were quite high, compared to the remainder of the group, making the distribution of duration thresholds markedly skewed for each word. To indicate the central tendencies of the thresholds, therefore, both means and medians were used. A third average threshold measure was derived by taking the mean of the ten lowest thresholds for each word. Means for these abridged distributions were almost identical with the medians. The time of exposure at which the threshold criterion was reached will, hereafter, be referred to as t; and t_M, t_{Md}, and t_L will, hereafter, stand for the three indices of central tendency for thresholds.

Experiment I

Table I presents alphabetically 60 words used in Experiment I. Another study (3) required that the list include 30 relatively common words—five representing each of six different value or interest categories—and 30 words of relative rarity, similar to, and preferably synonyms of, the common ones. Synonyms and words of the same value category were separated in the standard order of

[4] A word reported correctly upon its first exposure was assigned no threshold value, since it was impossible to estimate how short an exposure could have been reported correctly. Only 15 of the 1,200 thresholds in Experiment I and only 1 of the 300 thresholds in Experiment II had to be deleted for this defect. One of our 20 *S*s required exposures longer than 1,000 msec. for several of the words in Experiment I. Since the timer was not calibrated in that range, his data were discarded, leaving a total of 19 *S*s. In the rare cases (4 of 1,140) in which this problem arose with other *S*s, the threshold was arbitrarily recorded as 1,000 msec.

TABLE I

WORDS USED IN EXPERIMENT I

All letters were capitalized in Gothic type style

ALTRUISTIC	INTELLECTUAL
AMICABLE	JUDICIARY
ASSETS	KNOWLEDGE
ASSIDUOUS	LAWYER
AUTOMOBILE	LIBERTIES
BARRISTER	LIMOUSINE
BEATIFIC	LITERARY
BENIGN	MARKET
BROKER	METAPHOR
CELESTIAL	MUNDANE
CHANCELS	ORCHESTRA
CHARITABLE	PAINTING
CHURCHES	PEDAGOGUE
CONDOLENCE	PERCIPIENCE
CONVIVIALITY	PHYSICS
DEBATING	PICTURE
EARTHLY	POETRY
ECONOMICS	PSYCHICAL
EDUCATION	REBUTTAL
ELEGIES	RELIGIOUS
ENSEMBLE	REVERENCE
ERUDITION	SAVINGS
ETCHER	SCIENTIFIC
FRIENDLY	SERVICE
FRUGALITY	SPIRITUAL
GOVERNMENT	STATICS
HEAVENLY	SYMPATHY
HOSPITALITY	THEISTIC
INDUCTIVE	UNCOERCED
INITIATIVE	VIGNETTE

word presentation to minimize their interactions. The frequencies of these words in the three word counts (f_m, f_{sm}, and f_{av}) were converted into logarithms to facilitate their manipulation, since the range of frequencies was from zero to 3,000.[5]

[5] In the four counts of the *Teacher's Word Book*, words that occur with greater rarity than once per million words are not assigned separate frequencies for the different words, but the total times they occur in all four counts (totaling eighteen million words) is supplied. Divided by four, these frequencies are comparable to those for the more common words, except that no differences between the Magazine and Semantic Counts can be determined. Words that do not occur at all in the Thorndike-Lorge Counts, if assigned a value of zero, would have a log frequency of minus infinity. To make possible the use of such extremely rare words, we assumed that they would occur once on the average in a sample ten times as large as the Thorndike-Lorge word sample, thus giving them a log frequency of − 1.00.

RESULTS

Product-moment correlation coefficients relating the three measures of central tendency of threshold to the logarithms of the three measures of word-frequency are shown in Table II. The coefficients range from − .68 to − .75; covariance of log word-frequency with duration threshold accounts for an average of 50.9 per cent of the total variance. Of the three measures of word-frequency, f_{av} is more highly correlated with duration threshold than either f_m or f_{sm}, though these differences among the coefficients of correlation are not statistically reliable. Scatter-plots of the correlations (Fig. 1) justify the calculation of linear (product-moment) correlations between the frequency variable and the threshold variable.

TABLE II

PRODUCT-MOMENT CORRELATIONS BETWEEN LOG WORD-FREQUENCY AND DURATION THRESHOLD (N = 60) FOR EXPERIMENT I

Duration Threshold	Log Word-Frequency		
	log f_m	log f_{sm}	log f_{av}
t_M	−.68	−.72	−.73
t_{Md}	−.68	−.70	−.72
t_L	−.71	−.73	−.75

The data of Experiment I demonstrate a strong inverse relationship between relative word-frequency and duration threshold. Words of high frequency of occurrence require shorter exposure durations for correct report than do words of low relative frequency of occurrence.

EXPERIMENT II

Several uncontrolled factors were forced upon the previous experiment by the selection of words for a planned

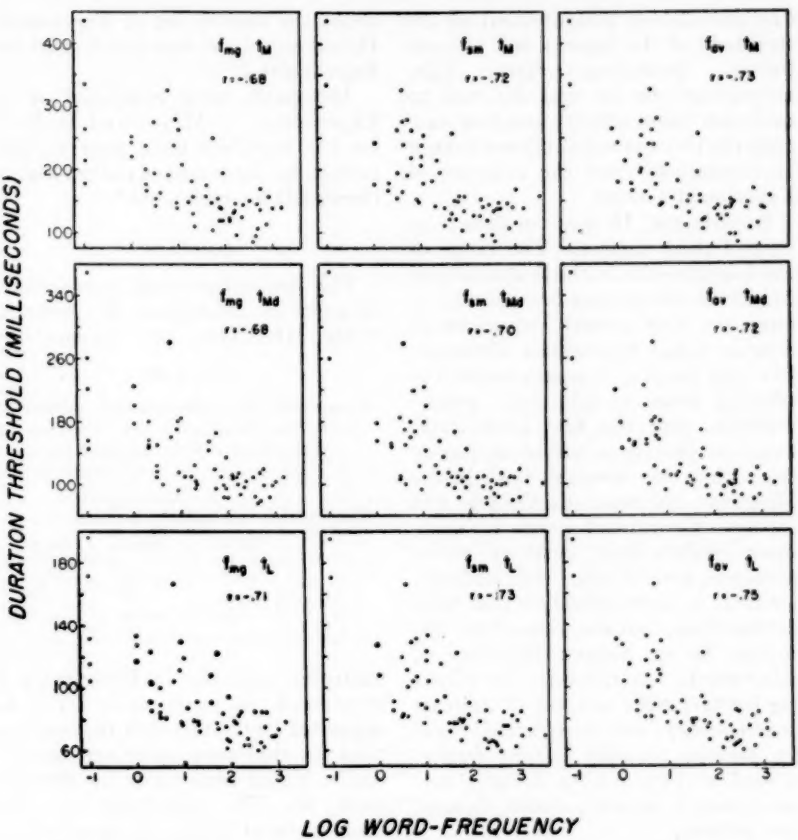

FIG. 1. Visual duration threshold as a function of word frequency. Scatterplots for the three measures of central tendency (means, medians, and means of the ten lowest thresholds) and three sources of word frequencies (Magazine Count, Semantic Count, and their geometric means).

subsequent experiment (3). The words varied from 6 to 12 letters in length. Plurals and *ing* forms, it has already been observed, cannot be distinguished from their roots in the published Thorndike-Lorge tables, and the frequencies for such words in Experiment I are probably inaccurate. For a few words, the frequencies based on the two counts are so disparate that even their ranks in the hierarchy of relative frequencies may be called into question. Strong associations, moreover, interrelate many of the words. For a pair of synonyms (or for words classified in the same value category), one following closely upon the other in the tachistoscopic list, the repetition of the first word during the measurement of its threshold might be expected to raise the probability (and lower the threshold) of the second. For example, the probability of *lawyer* might be raised by the repetition of

barrister, during measurement of the threshold of the latter a few seconds before. Scattering related pairs throughout the list was designed to minimize these effects; but how successfully this was accomplished cannot be determined from the evidence of Experiment I alone.

Experiment II was designed to control these factors. The range of word-frequencies in the Semantic and Magazine Counts was divided into 15 intervals, with central frequencies at roughly equal logarithmic distances. For each interval, a page number was selected from a table of random numbers, and the first seven-letter word on that page whose frequency lay within that interval, both in the Magazine and Semantic Counts, was chosen. Contextual words (prepositions, conjunctions, auxiliary verbs, pronouns, articles, etc.) were omitted, for there is some indication that their probabilities, outside connected discourse, do not behave like those of other words. In this way, the following list (arranged in order of decreasing frequency) was constructed: *country*, *promise*, *example*, *balance*, *deserve*, *venture*, *welfare*, *testify*, *dwindle*, *surmise*, *vulture*, *irksome*, *titular*, *figment*, and *machete*.

In random order, these 15 words were presented for threshold determination to the 20 *S*s, exactly as were the 60 words of the main list of Experiment I. Each *S* was given the words of Experiment II at a session separated by two weeks or more from Experiment I. Half of the *S*s were first given the 15-word list, and half received the 60-word list of Experiment I first. Despite the identical procedures, the comparative brevity of the 15-word list swelled the relative effect of practice on duration thresholds. Only two preliminary words (*HARVARD* and *TEACHER*) preceded the shorter list of Experiment II, instead of the four words used for Experiment I.

Thresholds were computed as in Experiment I. Means and medians for the complete data were supplemented by the means of the ten lowest thresholds for each word.[6]

Results

The product-moment correlations obtained in Experiment II appear in Table III. That the various uncontrolled variables in Experiment I introduced no important error is suggested by the fact that the correlations for that experiment are, on the whole, higher than those for Experiment II. The differences are too small to be of statistical significance. The average covariance of visual duration threshold with log word-frequency is only 44.5 per cent for Experiment II, compared with 50.9 per cent for Experiment I.

TABLE III

Product-Moment Correlations between Log Word-Frequency and Duration Threshold for Experiment II

Duration Threshold	Log Word-Frequency		
	log f_{ms}	log f_{sm}	log f_{sr}
t_M	−.71	−.79	−.76
t_{Md}	−.61	−.69	−.66
t_L	−.56	−.65	−.61

Differences between Experiments I and II

Although Experiment II yielded the same relation between duration

[6] None of the thresholds in this experiment was defective, but the data from one *S*, to whom the experiment was given under the wrong illumination, had to be discarded, leaving 19 *S*s.

TABLE IV

COMPARISON OF EXPERIMENTS I AND II WITH RESPECT TO SIZE OF DURATION THRESHOLD (in milliseconds) WHEN LOG f IS ZERO

Duration Threshold	Exper.	Log Word-Frequency		
		log f_{wg}	log f_{wm}	log f_{ww}
t_M	I	255	275	268
	II	321	333	329
t_{Md}	I	200	217	211
	II	243	252	249
t_L	I	125	134	131
	II	149	154	152

threshold and log word-frequency that was found in Experiment I, the two experiments give different absolute magnitudes of threshold for words of the same frequency. The central tendencies of thresholds measured in Experiment II are consistently higher than those for words of the same frequency in Experiment I. Table IV presents the duration thresholds for the two experiments at log $f = 0.00$. These values were obtained from least-squares fits of straight lines to the data summarized in Tables II and III. These differences between the magnitude of thresholds found in Experiments I and II will be examined here in the light of demonstrable differences between the procedures in the two experiments.

The principal difference between the two experiments is in the length of the list of words. Experiment II included only 15 words; Experiment I, 60. Otherwise the two experiments were matched as closely as possible. Since practice has a large effect upon magnitude of threshold, this difference between the two experiments is important. The effect of practice is shown in Fig. 2, in which the means of the 19 thresholds for each of the 60 serial positions in Experiment I (solid circles) and for

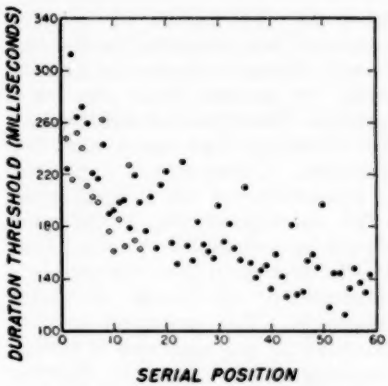

FIG. 2. Comparison of mean duration threshold for words of Experiments I and II as a function of serial position. Solid circles are means for Experiment I; open circles are means for Experiment II.

the 15 serial positions of Experiment II (open circles) are plotted. Over the period for which data are available for both experiments, the two practice curves are almost identical.[7]

Figure 2 shows that about three-fourths of the practice effect in Experiment I is accomplished in the first 15 threshold measurements, although they constituted only one-fourth of that experiment. Since in both Experiments I and II the serial position of each word was staggered throughout the list, about three-fourths of the threshold determinations for a word in Experiment I were made under conditions of practice beyond those attained in Experiment II. Thus it can be deduced from the difference in length of the experiments that the distribution of thresholds for a word in Experiment I will include more low thresholds than the distribution for a word of the same frequency in Experi-

[7] Actually the thresholds for Experiment II are slightly lower than those for Experiment I. This is accounted for by the fact that the mean log frequency of the words in Experiment II was higher.

ment II. The effect of practice, moreover, is to compress the distribution of duration thresholds; i.e., the effect is greater upon the third quartile of the threshold distributions, for example, than upon the first quartiles. Consequently, if practice is responsible for the difference between the experiments, the difference should be greater for the third than for the first quartile of the threshold distributions for words of stated frequency. This argument can be extended to any measures of central tendency of the threshold distributions. Thus the differences between Experiments I and II should be greatest for t_M, the measure of central tendency that depends most upon the upper end of the distribution of thresholds; less for t_{Md}, which depends only upon the central cases of the distribution; and least for t_L, which is affected only by the lower half of the distributions. The data, shown in Table IV, are in accord with the prediction.

On the basis of practice effects alone, therefore, the differences between Experiments I and II can be predicted. Consequently, we can assume that the various factors uncontrolled in Experiment I but controlled in Experiment II (length of word, associated interaction between synonym pairs, etc.) have only negligible effects upon the function relating duration threshold to word-frequency. Moreover, the fact that the differences between two experiments can be resolved with good accuracy lends credence to the generality of the relationship between duration threshold and log word-frequency. At present, the most likely form of this relationship appears to be given by the equation:

$$t = -k \log f,$$

where the slope constant k depends upon what measure of central tendency happens to be used to represent the distribution of thresholds.

Correction for Other Variables

A host of parameters, many introduced by characteristics of the apparatus, influence the absolute size of duration thresholds in the preceding experiments. Some of these were held constant by the procedure: illumination, size and style of type, and state of adaptation, for instance. But differences in the physical characteristics of the actual letters constituting different words cannot easily be eliminated by experimental design, although these factors must contribute a large portion of the error variance in the experiments above. Only actual experiments in which variables like word-frequency and practice are controlled and the physical characteristics of words are varied systematically can determine their precise effects on thresholds. Rough empirical estimates of some of them, however, can be obtained, when word frequency is held constant, from an examination of the thresholds of particular words in Experiment I that happen to feature such factors. We have been able to develop empirical corrections for such variables that strengthen the relation between word-frequency and duration threshold.

Six characteristics of the words in Experiment I were studied in relation to obtained visual duration thresholds. In general, it was found that, with frequency held constant, (1) letters composed of only a few continuous lines (*I*, *J*, *L*, *T*) tend to have low duration thresholds; (2) letters in which the white background is chopped up by numerous lines (*M*, *N*, *S*, *W*), or those in which the white back-

ground lies at the center of the letter so that its lines are closer to surrounding letters than to each other (*U*), or those whose striking similarity of form leads to conflicting responses (*C*, *G*, *O*), tend to have high thresholds; (3) words with average syllable lengths of less than 2.75 letters (which was the mean for all 60 words) tend to have lower thresholds than words made up of longer syllables; (4) words with repetitive patterns of letters, of the form *AA*, *ABA*, *ABAB*, tend to have low duration thresholds. To the extent that the small number of cases allowed, Ranschburg's observation (2) that the effects of letter repetitions vary inversely with the number of letters separating the repeated pair, appeared to hold.[8] The length of a word did not appear to affect our thresholds.

Corrections for the above four classes of variables which were thought to influence visual duration thresholds were incorporated into a formula with empirical weights assigned to each factor. When visual duration thresholds were thus corrected, the relationship between log word-frequency and visual duration threshold was increased. Product-moment correlations between corrected duration threshold and log word-frequency are presented in Table V. The efficacy of the corrections can be estimated by comparing the data of Table V with the data in Tables II and III. All correlations were increased by applying the empirical corrections. An average of 62.3 per cent of the variance in Experiment I can be accounted for by the covariance of corrected threshold with log word-frequency, a figure that is considerably higher than the 50.9 per cent observed for uncorrected thresholds. The corresponding increase in covariance for Experiment II is from 44.5 per cent to 55.3 per cent.

Although we do not want to attach any theoretical significance to an empirical correction procedure, it is suggestive that the same corrections have such a marked and identical effect upon data of both Experiments I and II. However, the legitimacy of the corrections must await further experimentation in which the important variables are systematically controlled. We should expect that such experiments would reveal a higher correlation between log word-frequency and visual duration threshold than we have been able to demonstrate in our experiments.

TABLE V

PRODUCT-MOMENT CORRELATIONS BETWEEN CORRECTED DURATION THRESHOLD AND LOG WORD-FREQUENCY FOR EXPERIMENTS I AND II

Duration Threshold	Log Word-Frequency		
	log f_{wg}	log f_{sm}	log f_{sv}
Experiment I			
t_M	−.76	−.76	−.79
t_{Ms}	−.79	−.76	−.81
t_L	−.82	−.78	−.83
Experiment II			
t_M	−.82	−.84	−.83
t_{Ms}	−.69	−.72	−.72
t_L	−.68	−.70	−.70

[8] The late Dr. Paul Schiller originally suggested that a correction be made for letter repetitions, and pointed out the relevance of Ranschburg's detailed study. Ranschburg felt that a significant proportion of errors in the reproduction of groups of letters were duplications or omissions of a repeated element. The observation here, that letter repetitions are associated with lower visual duration thresholds, can be made consistent with his results by assuming that response to one occurrence of a repeated letter greatly raises the probability that a response will be made to a second occurrence.

Summary

1. Data from two experiments, using 75 words, show that the visual duration threshold of a word, measured tachistoscopically by an ascending method of limits, is an approximately linear function of the logarithm of the relative frequency with which that word occurs in the Thorndike-Lorge word counts.

2. Product-moment correlations between the two variables range from − .68 to − .75 in the main experiment. Empirical corrections for certain physical characteristics of the words and of their component letters were found to raise the range of those correlations to − .76 to − .83.

3. Viewed as a correlation between language behavior in a highly specific situation (duration threshold determination) and in general usage (Thorndike-Lorge counts), these data are of significance for the experimental analysis of language behavior. In addition, they point to the necessity for operational interpretation of a large number of perceptual experiments, since the size of threshold is found to be a function of a property (relative frequency) of the responses (words) that operationally define perceptual threshold.

(Manuscript received April 26, 1950)

References

1. Lorge, I. The English semantic count. *Teach. Coll. Rec.*, 1937, **39**, 65–77.
2. Ranschburg, P. Ueber die Wechselwirkungen gleichzeitiger Reize im Nervensystem und in der Seele. *Z. Psychol. physiol. Sinnesorg.*, 1913, **66**, 161–248; 1914, **67**, 22–144.
3. Solomon, R. L., and Howes, D. H. Word-probability, personal values, and visual duration thresholds. *Psychol. Rev.*, 1951, **58**, 256–270.
4. Thorndike, E. L., and Lorge, I. *The teacher's word book of 30,000 words.* New York: Teachers Col. Bur. Publ., Columbia Univ., 1944.

THE EFFECT OF NEGATIVE INCENTIVES IN SERIAL LEARNING: VI. RESPONSE REPETITION AS A FUNCTION OF AN ISOLATED ELECTRIC SHOCK PUNISHMENT

BY G. RAYMOND STONE AND NORMAN WALTER[1]

University of Oklahoma

The problem of the present paper stems directly from an earlier experiment by the senior author (**14**) in which it was demonstrated that the effect of an isolated verbal punishment in a serial task was *not* to weaken the strength of the stimulus-response relationship to which it was applied. The probability of the occurrence of the punished response was, instead, increased. Together with other data gathered in a coordinated research program (**12, 13, 14, 15, 16**), these results lead to the hypothesis that the effect of punishment in perceptually difficult serial tasks when the situation is not complicated by immediately available responses as alternatives to the one just punished (i.e., the effect of punishment per se) is fixative.

This hypothesis is consonant with, but not the same as, the viewpoints of Thorndike (**18**) and Skinner (**9**), both of which, while agreeing that punishment does not weaken, give to punishment the effect of inducing alternative responses (**3**). According to the present formulation, alternative responses occur more frequently than the punished ones only when they are rewarded or escape the punishment or achieve lesser punishment.

The present research investigates the influence of an informative serially isolated punishment, conveyed by an electric shock, upon the selection and fixation of verbal responses. The experimental design is identical with that of Stone's 1948 study (**14**) except for the employment of the stronger type of punishment.

Procedure and Apparatus

Two groups of 80 *S*s each performed in a serial verbal multiple-choice task in which a stimulus items (three-letter words) were presented at a 3-sec. rate by a Baker-Stone memory drum (**1**). Each *S* was read the following instructions:

> You are to be presented a list of words exposed one at a time in the opening now covered by the black tape. To each word in this list some one number or group of numbers between one and ten has been arbitrarily assigned in a random fashion by the experimenter. You are to respond to each word immediately with the first number between one and ten, inclusive, which occurs to you. After each of your responses I will do one of three things: (1) If the number you call out is one of the ones that has been assigned to that word, I may answer "Right"; (2) if your response is not one of the ones assigned to that word, I may say "Wrong" or administer an electric shock to your wrist; or (3) I may say nothing at all, in which case you will not know whether you are right or wrong.
>
> We will go through the list several times; there will be only three seconds between presentations of the words, and you must respond quickly. The first time we go through the list, of course, there is nothing for you to do but call numbers to the words at random, because the only way you can learn which numbers are correct is in terms of my responses. If your response to some word has been called "right," that response, if you can remember it, will always be right; the same is true for responses called

[1] The authors are indebted to the Faculty Research Committee of the University of Oklahoma for the costs of the current-stabilized shocking equipment, including the labor cost for Mr. Thaddeus J. Baker, who is responsible for the apparatus construction and to whom, therefore, the authors are also indebted.

"wrong" or shocked; that is, these responses, if repeated for that word, will always be wrong. In the cases where I say nothing, you will never know whether your responses are right or wrong for that word. There are forty words in the list. There is no logical or systematic arrangement of correct responses to help you. Instead, I repeat, the only cue you will have will be my responses of "right" or "wrong" or shock. Any attempts on your part to find a systematic order of right responses will decrease your efficiency. There is no system. Remembering the cues I call out and responding accordingly is the only way you can learn.

It is obvious that under these conditions you cannot possibly learn the entire list completely, and, as a matter of fact, we will stop the learning before you even approach completion. But we *are* interested in your percentage of correct responses and the degree to which you improve; i.e., repeat the responses called right and vary those called wrong or shocked until you have the right numbers. It is possible that more than one number may be correct for any one word, and there is always at least one number between one and ten that is correct. Some word-number combinations will be more difficult to learn than others, therefore, and you will undoubtedly still be missing some when we stop. But try to get as many as you can.

Be sure that you call some number for every word, and be sure that you respond quickly so that the three-second rate per word can be maintained.

Any questions raised by *S* were answered by re-reading that part of the instructions pertaining to the question or by stating an answer as nearly in keeping with the instructions as *E* was able. As soon as it appeared that *S* understood the instructions, *E* said, "You should now remove the tape and give me a number for the word you see." With the first answer the motor of the drum was started and the words automatically appeared at the 3-sec. rate per item. Ten trials were run with no intertrial rests. Each of 40 *S*s started with a different word. Thus the first word seen by the first *S* was *DYE*, and the second, *WAR*, and the third, *ARM*, etc. The second *S* would then be presented with the word *WAR* first, and then *ARM*, etc., the fortieth word being *DYE*. With 80 *S*s in each group, each stimulus item occupied every position twice for each condition.

Unknown to the *S*, the *E*'s informative incentive responses were given in a fixed order. The following symbols were used:

N—no response by *E* to *S*'s response
R—the word "right"
W—the word "wrong"
S—current-stabilized electric shock applied to *S*'s right wrist, indicating a wrong response
X—the variable incentive position; S in the experimental group, N in the control group

The order of *E*'s responses was: NWRWNRR-SNWWRNRWNN*NNNXNNN*NNWRWWN-NRNRSRWNN.

Data are analyzed only for the central incentive series (italicized above) located at what is considered the approximate middle of the list in order to minimize any possible general intraserial effects. The experimental group will be identified in the tables as performing under the NSN condition, the controls under the NNN(S) condition.[2]

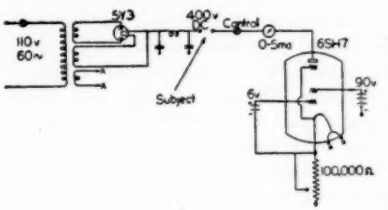

FIG. 1. Circuit diagram of the stabilized constant-current type shocking equipment

The Baker-Stone memory drum is equipped with a plywood screen attached to its front end which effectively blocks off *S*'s vision from *E* and the electric shock equipment other than the the electrodes. The circuit for the shocking apparatus is given in Fig. 1. A shock intensity of 1.67 mamp. current was used throughout.

RESULTS AND DISCUSSION

All data analyses are in terms of the number of times the *S*s repeated their responses to a list item as a function of the incentive response to that item. Serial-position analyses for the central incentive series are made at three

[2] The NNN(S) designation is used to distinguish the present data where shock was delivered to responses outside the central incentive series from comparable data of other studies (**14, 15**) where no shock was delivered anywhere in the entire series.

stages of practice: trial 2, indicating the very first effects of the incentive series; trials 2–6, a comparative measure for other studies of like nature; and trials 2–10, to demonstrate the changes in late stages of practice. A specific practice analysis is also made of the repetitions per trial to the X item only, i.e., the shocked item, X(S), of the experimental group compared to its comparable empirical control item, X(N), in the control group.

For any one *S* the possible range of response repetitions for any one item in six trials is from zero to five. The distribution of repetitions for all *S*s of one condition to one item shows severe skewness in the direction of the zero frequency, as is shown in Table I. To estimate the reliability of mean differences, therefore, a transformation of the raw data was in order. As suggested by Snedecor (11, p. 446) for data of this type, the transformation factor $\sqrt{X + 0.5}$ has been employed whenever necessary.

TABLE I

FREQUENCY DISTRIBUTIONS OF THE REPETITION OF RESPONSES IN TRIALS 2–6 AS A FUNCTION OF THE CENTRAL INCENTIVE SERIES
($N = 80$)

Central Incentive Series	Condition	Range of Repetitions					
		0	1	2	3	4	5
N	NSN	31	23	17	6	1	2
	NNN(S)	24	33	14	7	1	1
N	NSN	30	29	15	3	2	1
	NNN(S)	30	30	11	6	2	1
N	NSN	33	25	17	2	2	1
	NNN(S)	39	24	6	4	5	2
X	NSN	25	31	11	8	5	0
	NNN(S)	29	26	13	6	4	2
N	NSN	19	26	21	8	4	2
	NNN(S)	28	31	14	4	2	1
N	NSN	27	28	14	5	4	2
	NNN(S)	30	24	14	9	3	0
N	NSN	35	20	15	6	2	2
	NNN(S)	31	26	11	5	5	2

Serial-position analyses.—The serial position data are presented in Table II. The following generalizations describe the results:

1. Mean repetition to the shocked S-R shows no significant difference from mean repetition to its comparable control S-R which received no shock.

2. There is a relative tendency for repetitions to the shocked S-R to occur less frequently with practice than repetitions to its control S-R. This effect is more clearly demonstrated by the practice analysis to be presented later.

3. There is slight evidence (nonsignificant and nonsystematic) for the development of induced variability (non-repetition) in the S-R's preceding the shocked item as a function of practice.

4. There is a failure of shock to induce variability in the S-R's immediately following its administration. Instead, there is a significant increase of repetitions to the first S-R following shock, which in turn is followed by a decreasing gradient of repetitions for the rest of the series.

The data for the shocked item fail either to confirm or reject the hypothesis stated earlier. By this analysis there is no evidence for either a strengthening or a weakening effect attributable to shock. Further discussion of this point will be delayed until the presentation of the practice analysis.

The "spread" effects of the shock to the S-R's preceding and following it are problematic. The bidirectional gradient of induced variability from punishment (7) receives no support.

TABLE II

REPETITION OF RESPONSES AS A FUNCTION OF THE CENTRAL INCENTIVE SERIES
(N = 80 EACH GROUP)

A. Data from Trial 2

Condition	Measure	Central Incentive Series						
		N	*N*	*N*	*X*	*N*	*N*	*N*
NSN	Total	20	15	10	11	21	13	17
	Per cent	25.0	18.8	12.5	13.8	26.3	16.3	21.3
NNN(S)	Total	19	8	8	10	13	17	19
	Per cent	23.8	10.0	10.0	12.5	16.3	21.3	23.8
% Diff. NSN–NNN(S)		1.2	8.8	2.5	1.3	10.0	−5.0	−2.5

B. Data from Trials 2–6

Central Incentive Series	Per Cent Scores				% Diff.	$\sqrt{X+0.5}$ Transformations	
	NSN		NNN(S)				
	Total	Per Cent	Total	Per Cent		Mean Diff.	SE Diff.
N	89	22.3	91	22.8	−0.5	.0260	.0663
N	81	20.3	83	20.8	−0.5	.0064	.0671
N	78	19.5	78	19.5	0.0	.0240	.0819
X	97	24.3	96	24.0	0.3	.0163	.0721
N	118	29.5	84	21.0	8.5	.1582*	.0686
N	97	24.3	91	22.8	1.5	.0263	.0721
N	86	21.5	93	23.3	−1.8	.0336	.0755

C. Data from Trials 2–10

N	172	23.9	182	25.3	−1.4	.0419	.0959
N	179	24.9	174	24.2	0.7	.0404	.0938
N	157	21.8	199	27.6	−5.8	.1200	.0970
X	179	24.9	203	28.2	−3.3	.0566	.0889
N	206	28.6	184	25.6	3.0	.0900	.0883
N	184	25.6	191	26.5	−0.9	.0227	.0975
N	181	25.1	204	28.3	−3.2	.0703	.0954

* Only difference significant at the .05 level or better.

The reversed gradient in the post-punishment series, which completely violates expectations from that hypothesis, is a characteristic phenomenon in studies of this nature (**12, 13**).

A claim has recently been made by Smith (**10**) that a "probability bias" hypothesis can account for the spread of variability, and although this explanation seems credible for the results of Muenzinger and Dove (**8**), it is not clear how it would apply when the postpunishment gradient is reversed as in the present and other studies. According to this hypothesis, a spread-of-reward effect is accounted for by the nonrandom response tendencies (refractoriness to serial repetition) that follow a trial-to-trial repeated (correct) response, and the spread-of-punishment effect by

the nonrandom response tendencies that follow a nonrepeated (incorrect) response. In order for the hypothesis to apply to the present data, therefore, punishment must decrease repetitions to the S-R it follows, and it has been demonstrated that it does not do so. Even if it did, however, the reversed nature of the gradient as well as the significant absolute increase of repetitions to the first S-R following the shock would violate predictions from the hypothesis.

It has been suggested previously (13) that the gradient effect following punishment might be due to an indirect fixative action of the punishment itself. The process involved might be a function of rapid changes in motivational level, or its might be related to previous learning habits, or both. It is probable that each *S* enters the experimental situation with some number preference which he may use as a first response more consistently after the disruption which accompanies a punished item than at other times in the series. This preference would represent a more "available" response, to use Thorndike's term (17), and its use immediately following the punishment might involve behavior which has been described as "regressive" in the sense that it represents a reversion to an easier, more basic, or more primitive response. This account bears some similarity to the "rehearsal-distraction" hypothesis offered by Jenkins and Sheffield (6) to account for the data of spread-of-reward effect.

Most hypothetico-explanatory accounts of the spread-of-reward effect endeavor to use the same principle(s) to account for the pre- and postreward gradients, and a similar tendency shows itself in accounts of the spread effects of punishment; whereas the data in the latter case strongly suggest differential principles (13). Pre- and postpunishment effects appear to be opposed.

Practice analysis.—The probability that *S* will repeat on the second trial the exact response he gave for the same stimulus item on the first trial is 10 per cent by chance alone insofar as he has only ten available responses. To expect *Ss* to repeat responses at this percentage in the absence of incentive information (the N conditions), however, assumes that the *Ss*' original response tendencies were of equal strength for all ten available responses, whereas it is definitely established (2, 5, 10) that *Ss* do not respond with randomness in this situation. It is safe to assume, therefore, that some original biases from number preferences existed. Furthermore, it seems probable that the *Ss* might attempt to relate meaningfully a stimulus item to one of their available responses in the event that, if their response would be called "right" they could remember it to repeat it, or if called "wrong" they could remember it so as not to repeat it. If, however, their response is followed by no information, there still will exist an association bias that might well have the function of producing for *S* an increased perceptual clarity. These two biases (original number preferences, and the *Ss*' "provisional tries") would operate to increase repetition on Trial 2 above chance expectancy even when no information is given, and such, as a matter of fact, is the case in the control (NNN) condition. For the X(N) item in the middle of the central incentive series, for example, the per cent repetition on Trial 2 is 12.5. For the X item of the NSN central series, where in addition to these biases the *Ss*' Trial 1 responses received the informative electric shock, one might expect, on

TABLE III

REPETITION OF RESPONSES AT X AS A FUNCTION OF PRACTICE
(N = 80 EACH GROUP)

Condition	Measure	Trials								
		2	3	4	5	6	7	8	9	10
NSN	Total	11	18	24	24	20	19	22	16	24
	%	13.8	22.5	30.0	30.0	25.0	23.8	27.5	20.0	30.0
NNN(S)	Total	10	16	27	19	24	23	29	27	28
	%	12.5	20.0	33.8	23.8	30.0	28.8	36.3	33.8	35.0
% Diff. X(S)–X(N)		1.3	2.5	−3.8	6.2	−5.0	−5.0	−8.8	−3.8	−5.0

the "provisional try" hypothesis (**4**), fewer repetitions than this. The data for this item, however, show 13.8 per cent repetition. That the shock had no weakening effect on the S-R relationship to which it was applied receives support, but that this relatively small difference of response repetition in favor of the shocked response could be considered evidence that the S-R had been strengthened by the shock is still uncertain.

Beginning with the third trial a new bias complicates the analysis. Because of the instructions to repeat as many right responses as possible (but also to vary the wrongs), the *S*s may have possessed a generalized motive to repeat responses. If, therefore, they repeat a response on Trial 2, with perceptual recognition of the repetition, this should act as a subjective positive incentive and thus increase repetitions on later trials. If this tendency were operative, we should expect the per cent repetition to the X(N) item of the control group to show additional increases with practice, and such, as a matter of fact, is the case, as can be seen in the data presentation of Table III. For the X(S) item of the experimental group, however, a perceptual recognition of a repetition of a response which was previously shocked and is again shocked could hardly act as a subjective positive incentive. If the effect of the shock is supposed to be weakening, one would expect the repetitions to the X(S) item to show a decrease with practice. That the reverse, actually, is the case (see Table III) seems to the writers to be evidence for a strengthening effect of shock. At any rate, some strengthening effect is operating, although less in the later stages of practice (see Fig. 2) than the strengthening operating in the control condition.

To say, as some might, that the *S*s' responses to the X(S) item were subject to both the strengthening action that operated for X(N) and a weakening action of the shock, and that this accounts for the late practice

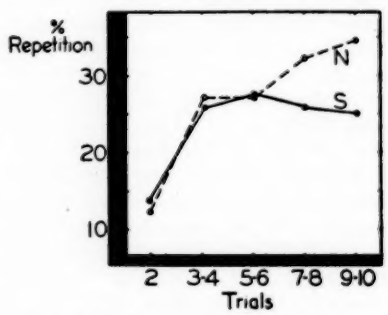

FIG. 2. Per cent response repetition to the X item as a function of practice

differential in favor of the X(N) condition, appears a little far-fetched. This point of view would have to explain how a perceptually recognized repeated shocked response which was itself shocked could act as a subjective positive incentive. The *S*s themselves frequently reported that they knew the response they were making *had been* wrong and *had been* shocked, that it *was* wrong and *was going to receive* shock, but they "made it anyway" and were very disgusted with themselves for doing so, and wanted to know *why* they did it. The *E*s told them that the shock strengthened their tendency to respond. An answer in terms of "expectancy" theory (**19**) did not seem to apply.

Summary

The hypothesis that the effect of punishment per se (i.e., not complicated by the effects of immediate alternative responses which may achieve reward, no punishment, or lesser punishment) is fixative has been investigated in a serial verbal multiple-choice situation. Two groups of 80 *S*s each gave data to compare the selective effects of a serially isolated electric shock punishment with a comparable empirical control condition where no incentive was present.

From a practice analysis of the punished S-R over ten trials, the hypothesis received support. Additional analyses of the spread-of-punishment effect to adjoining non-incentive S-R's gave complicated data which were differential as a function of pre- or postpunishment position. In the postpunishment series there was some evidence for an induced fixation with gradient characteristics. The data are not easily explained by hypotheses which have been offered by Muenzinger (**7**) and Smith (**10**). An alternative was suggested.

(Manuscript received May 17, 1950)

References

1. Baker, T. J., and Stone, G. R. Design and construction of a research memory drum. *Proc. Okla. Acad. Sci.*, 1948, **29**, 108–113.
2. Eagleson, O. W., and Lipford, E. J. A study of number choices. *J. gen. Psychol.*, 1944, **31**, 129–133.
3. Estes, W. K. An experimental study of punishment. *Psychol. Monogr.*, 1944, **57**, No. 3.
4. Hilgard, E. R. *Theories of learning.* New York: Appleton-Century-Crofts, 1949.
5. Hull, C. L. Special review of Thorndike's *Fundamental of human learning. Psychol. Bull.*, 1935, **32**, 807–823.
6. Jenkins, W. O., and Sheffield, F. D. Rehearsal and guessing habits as sources of the "spread of effect." *J. exp. Psychol.*, 1946, **36**, 316–330.
7. Muenzinger, K. F. The law of effect: a round table discussion. *Psychol. Rev.*, 1938, **45**, 215–218.
8. Muenzinger, K. F., and Dove, C. C. Serial learning: I. Gradients of uniformity and variability produced by success and failure of single responses. *J. gen. Psychol.*, 1937, **16**, 403–413.
9. Skinner, B. F. *The behavior of organisms.* New York: D. Appleton-Century, 1938.
10. Smith, M. H. Spread of effect is the spurious result of non-random response tendencies. *J. exp. Psychol.*, 1949, **39**, 355–368.
11. Snedecor, G. W. *Statistical methods.* Ames, Ia.: Iowa State College Press, 1946.
12. Stone, G. R. The effect of negative incentives in serial learning. I. The spread of variability under electric shock. *J. exp. Psychol.*, 1946, **36**, 137–142.
13. Stone, G. R. The effect of negative incentives in serial learning. II. Incentive intensity and response variability. *J. gen. Psychol.*, 1950, **42**, 179–224.
14. Stone, G. R. The effect of negative incentives in serial learning. III. Fixation due to an isolated verbal punishment. *J. gen. Psychol.*, 1948, **38**, 207–216.
15. Stone, G. R. The effect of negative incentives in serial learning. V. Response repetition as a function of successive serial verbal punishments. *Amer. Psychologist*, 1949, **4**, 227 (Abstract).

16. Stone, G. R., and Yerrington, F. The effect of various patterns of verbal incentives upon response repetition. *Proc. Okla. Acad. Sci.*, 1948, **29**, 129–131.
17. Thorndike, E. L. *Selected writings from a connectionist's psychology*. New York: Appleton-Century-Crofts, 1949.
18. Thorndike, E. L. *The fundamentals of learning*. New York: Teacher's Coll., Columbia Univ., 1932.
19. Tolman, E. C. *Purposive behavior in animals and men*. New York: D. Appleton-Century, 1932.
20. Walter, N. The specific and spread effects of an isolated electric shock punishment in a serial verbal multiple-choice situation. Unpublished Master's thesis, Univ. of Oklahoma, 1949.

PAIN MEASUREMENT BY THE RADIANT HEAT METHOD: INDIVIDUAL DIFFERENCES IN PAIN SENSITIVITY, THE EFFECTS OF SKIN TEMPERATURE, AND STIMULUS DURATION

BY JAMES E. BIRREN, ROLAND C. CASPERSON, AND JACK BOTWINICK

National Heart Institute, National Institutes of Health, and Gerontology Section, Baltimore City Hospitals

Studies of pain always have been difficult because of the problems of obtaining a stimulus for pain which can be accurately controlled and measured. In recent years a method (**6**) employing a graduated source of heat has been used to measure the analgesic properties of drugs (**1, 10, 11, 13, 14**) and also as a source of pain in studies of the responses of neurotic individuals (**4, 5, 9**). The radiant heat apparatus of Hardy, Wolff, and Goodell (**6, 7**) focuses a large incandescent lamp upon a small area of the forehead that has been blackened with India ink to insure complete absorption of the radiation. The amount of heat can be varied by changing the voltage to the lamp. By appropriately varying the heat stimulus, measurements of the pain threshold may be obtained. Subjectively there is little difficulty in identifying as pain the sensation produced by an intense radiant heat stimulus.

Because of the control that can be exercised over the pain-stimulus, the radiant heat apparatus was used in an attempt to explore age differences in salicylate analgesia (**2**). Since this study found not only a lack of an age difference but also a lack of any analgesia (also **8**), it was decided to investigate the technique from the point of view of reliability of measurement.

Although the radiant heat method of pain measurement has been well studied from its physical aspects, its reliability in a psychophysical sense has not been given serious attention. The range of individual differences in the threshold stimulus has been expressed as a percentage of the mean (**6**). The use of such a percentage expressing the ratio of the range of individual differences divided by the measured stimulus value is of doubtful validity. One could reduce this percentage by simply inserting a series of absorption screens in front of the skin. It is very likely that only a small portion of the energy of the lamp stimulates the nerve endings and that the major portion of the heat is absorbed by inert tissues, i.e., inert in this context. Thus the small percentage expressing variability could imply that the technique is poorly adapted to the measurement of individual differences in pain sensitivity. A suitable evaluation of the technique should include determinations of its reliability with respect to the range of individual differences.

These considerations prompted the present study whose purposes were to determine: (1) the reliability of pain measurements by the radiant heat method as might be expected in routine measurements, using the general technique of the previous investigators, (2) the day-to-day variability in pain sensitivity, (3) the effects of skin temperature on pain measurements, (4) the rate of dissipation of the heat that is imparted to the skin

by a single pain stimulus, and (5) the relation of stimulus intensity to exposure time when skin temperature is held constant.

Procedure

Apparatus.—The Hardy, Wolff, and Goodell apparatus was used in all pain measurements (6, 7).[1] A voltage regulator minimized fluctuations in the a.c. supply to the pain apparatus. Control of head position was aided by the use of a chin rest.

Determinations of the radiant energy from the lamp were made with a calibrated radiometer.[1] Calibration of the shutter was made with a Cambridge slit camera; measurements were made to .1 sec. Shutter times were constant within this ± limit but differed from the scale reading as follows: 3 sec., 0.0; 5 sec., + .1; 8 sec., + .4; 10 sec., + .6. For skin temperature measurements a Rauh Surface Pyrometer was used.[2] Measurements were made to .2 deg. Fahrenheit. Calibration of the pyrometer was made against a water bath of known temperature determined by a calibrated mercury thermometer (error ± .01° F.). Pyrometer measurements were linearly related to the calibrated temperature in the range 77° to 105° F. with a range of error of + .1 to − .3; this error is close to the limits of accuracy of scale reading, i.e., .2° F.

Measurements.—The forehead was washed with mild soap, dried, and wiped with a small amount of 95 per cent ethyl alcohol. Several coats of India ink were then applied, blackening an area of about 1¼ in. in diameter. Several coats of India ink were used to insure absorption of the radiant energy. Additional ink was applied whenever flaking occurred.

Ss were given detailed instructions before their initial pain measurement. They were told they had to distinguish between heat and the drawing pinprick sensation of pain. The sensation was described to *S* as a feeling of heat building up to a level where the skin seemed to feel as though it were drawing together, with a final sharp stinging sensation.

In determining the reliability of the technique, the general method of Hardy *et al.* was used (6). The number of stimuli per series was increased, however, to include pain and non-pain stimuli before beginning an ascending series. Between two and five stimuli were used to acquaint *S* with the distinction between heat and pain. Within the range described by the pain and non-pain stimuli, an ascending series of stimuli was imposed. An increment of 10 millicalories was used and about six stimuli (range 4 to 15) were used per *S*. The pain measurement was defined in accord with the previous investigators (6) as the lowest stimulus which was reported as pain. At least two positive reports were obtained in each series. In the skin temperature and stimulus duration studies a method of minimal changes employing stimuli randomly above and below the threshold was used.

To control skin temperature, a cool and a warm water bag with dry surfaces were used. The procedure consisted of either cooling or warming the forehead to a few degrees beyond the desired level. Serial skin-temperature measurements were then made until the skin reached the desired temperature. Within 5 sec. after the skin temperature reached the proper level, the inked area of the forehead of *S* was placed against the aperture of the apparatus and exposed to the stimulus lamp.

To determine the optimum time interval between pain stimuli, serial skin temperature measurements were made after a single high intensity stimulation. Three *Ss* were given a 3-sec. 225-millicalorie stimulus at a skin temperature of 90° F. Dissipation of the heat imparted to the skin followed a uniform time course (Fig. 1). The mean value of the three *Ss* show an asymptotic level at about 2 min. Thus, for any experiment in which *S*'s forehead is allowed to cool at room temperature, an interval of 2 min. or more is indicated. An interval of this duration would avoid accumulation of heat in the skin which would affect the determination of the threshold level. In practice, the method of alternating two *Ss* during

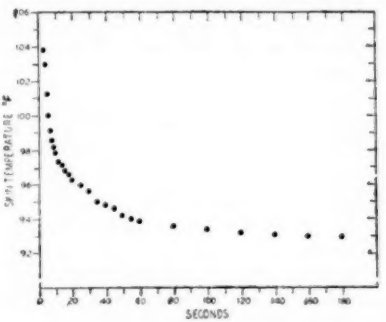

FIG. 1. Skin temperature of the forehead following a single exposure to a painful stimulus of 3-sec. duration at 225 millicalories/cm². Initial skin temperature was 90° F.

[1] Experimental Engineering Corporation, 96 South Washington, Bergenfield, New Jersey.

[2] E. M. Rauh, Inc., Buffalo, New York.

the measurements is convenient since it avoids the necessity of waiting while *S*'s skin cools.

Subjects.—The *Ss* for the reliability study were 46 professional and technical personnel of the gerontology laboratory at The Baltimore City Hospitals and 4 residents of the Old People's Home of the hospital. Of this group 30 were male, 20 female; 47 were white and 3 colored. The age range was 21 to 64 yr., with a median of 28 yr. Nine and 10 *Ss*, respectively, were used in the temperature and stimulus duration studies. These were all laboratory personnel highly experienced on the pain apparatus.

Results

Reliability.—For each of the 50 *Ss* two pain measurements were made in a single experimental session. To complete an experimental session required about 30 to 60 min. for two *Ss*, depending upon the number of stimuli necessary. The correlation between measurement 1 and 2 was .91, yielding an error of measurement of about 12.5 millicalories. When the measurements were repeated on the next day, the correlation between 3 and 4 was .86, with a corresponding error of measurement of about 15.4 millicalories.

There was no significant change in the mean level between the first and second determinations on a single day or between Day 1 and Day 2 (Table I), indicating a lack of a learning or experience effect.

TABLE I

Mean Pain Measurements of 50 Ss Determined Twice on Two Successive Days

Values are given in millicalories/cm² for a 3-sec. exposure.

	Day 1 Determination		Day 2 Determination	
	A	B	A	B
Mean	184.2	180.2	186.5	183.0
σ	30.2	29.0	30.6	32.0

For routine measurements, a mean of two determinations of the pain threshold was accepted by the present investigators as an adequate procedure.

Day-to-day variability.—The variability in pain sensitivity between two successive days was determined by correlating the mean of the two determinations for each *S* on each of the two successive days. The Day 1 vs. Day 2 correlation was .85 ($N = 50$). If one had a perfectly reliable measure of pain, the correlation between Day 1 and Day 2, corrected for attenuation, would be about .91. There would thus be between 10 and 15 millicalories day-to-day variability even if one assumes an infallible measure.

Effects of skin temperature.—The effects of skin temperature upon pain measurements were studied at three temperature levels: 85°, 90°, and 95° F. Nine *Ss* had two threshold determinations at each temperature. Only one temperature level was used per *S* per day. *Ss* were randomized in the order of exposure to the different temperature levels.

The three levels of skin temperature resulted in statistically significant changes in pain threshold. The nature of the distributions at each temperature can be seen in Fig. 2.

In order to interpret the significance of skin temperature as a factor in determining the range of individual differences in pain thresholds, a measurement of the distribution of skin temperatures under the general conditions of this study was needed. For this purpose, determinations of the temperature of the forehead in three areas contiguous to the region used for pain measurements, were made on 50 *Ss*. Measurements were made during a single work morning when the dry air temperature was

76° F. The mean skin temperature was 92.8° ($\sigma = 1.1$). This mean value is close to that accepted by Hardy *et al.* as the constant value, i.e., 93.2° F. The correlation between the first two areas measured on each *S* was .85; the estimated reliability of the mean value based on three measurements was .94, with a standard error of measurement of .38° F.

The absolute range of temperature values was 90.3° to 94.8° F. With a range of individual differences in skin temperature of approximately 5.0° F., even when environmental temperature was approximately constant, a nonsensory influence on the pain threshold has been identified.

Stimulus duration.—Since skin temperature was found to have a significant effect upon the pain threshold, the effects of variations in stimulus duration were studied when the initial skin temperature was held constant (90° F.). Varying the stimulus duration produced significant changes in threshold level. As seen in Fig. 3, there was no overlap in the distribution between the 3- and 5-sec. exposure. It is reasonable to expect that the relation of the threshold to stimulus duration is of the type, Threshold = $K \log t + a$. The relation of the constants of such an equation to initial skin temperature is a problem for further research.

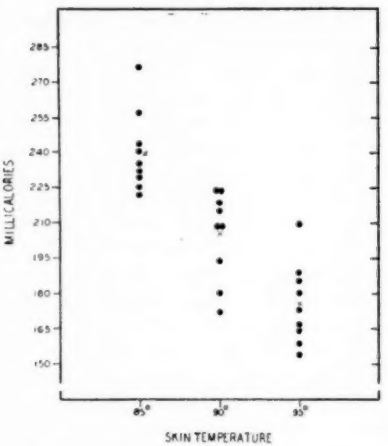

FIG. 2. Effects of initial skin temperature upon the pain threshold. Temperatures are given in degrees Fahrenheit, thresholds in millicalories/cm². X = mean threshold.

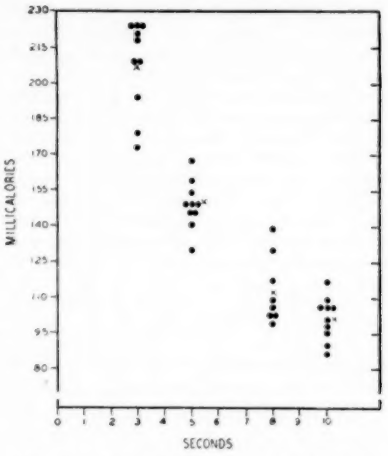

FIG. 3. Pain thresholds at stimulus durations of 3, 5, 8, and 10 sec. Initial skin temperature was 90° F. Thresholds are given in millicalories/cm². X = mean threshold.

DISCUSSION

The principal advantage of the radiant heat method of measuring pain lies in the fact that the stimulus is easily controlled and measured. In addition, the perception of a difference between heat and pain appears to be a discrimination that most young adults can make. A disadvantage of the method lies in the fact that an occasional *S* will not report pain until he has reached a stimulus level sufficient to result in a blistering of the skin. In addition to the occasional blister which might result from the method is the problem of con-

trolling skin temperature. The present study indicated that experimentally induced changes in skin temperature have an effect on the pain threshold as measured by this technique.

There are three sources of variations in skin temperature to be considered in connection with the radiant heat method of measuring pain thresholds: (1) fluctuations induced by variations in the environmental temperature, (2) individual differences in basal skin temperature, and (3) diurnal variations in skin temperatures.

Increase of the environmental temperature from 68° to 90° F. may produce an increase in skin temperature of the forehead of 2.5° F. in some *Ss* (**12**). In *Ss* kept for 1 hr. at a constant temperature (77° F.) individual differences in skin temperature may amount to 1.3° F. (**12**). Diurnal fluctuations under conditions of constant environmental temperature may be as large as 2.7° F. (**12**). When such factors are allowed to vary in random fashion, it can be appreciated that the pain threshold is influenced by nonsensory determinants when the thresholds are measured by the radiant heat method.

If an accurate assay of the efficacy of an analgesic is desired, skin temperature should be held constant to exclude influences of the drug which might alter body temperature and/or cutaneous blood flow. Such influences would alter the measurement of the peripheral thresholds and yet leave the mechanism of pain perception unchanged. Such reservations also apply to the interpretation of pain measurements on aged individuals. If individual differences in skin temperature are to be excluded from threshold measurements, it is likely that the number of determinations necessary to achieve satisfactory measurements will be larger than that employed by either the present or prior studies because of the constricted range of values.

Summary and Conclusions

1. A study was made of the radiant heat method of Hardy, Wolff, and Goodell in measuring pain sensitivity. The purpose of the study was to determine the reliability of the method, the day-to-day variability in the pain threshold, and the influence of skin temperature and stimulus duration.

2. Determinations of the rate of dissipation of the heat imparted to the skin by the pain stimulus indicated that at least 2 min. of rest must elapse between successive stimuli applied to *S*.

3. In routine pain measurements a correlation of .91 was obtained between two successive determinations on 50 *Ss*, with a standard error of measurement of 12.5 millicalories. The correlation between the mean of two determinations on Days 1 and 2 was .86 for the same 50 *Ss*.

4. When stimulus duration was held constant at 3 sec., a significant variation in pain threshold was produced by experimentally varying the skin temperature at levels of 85°, 90°, and 95° F.

5. When the skin temperature was held constant at 90° F., a significant relation was found between threshold and stimulus duration.

6. The radiant heat method of pain measurement offers the advantage of an easily controlled and measured stimulus. In any situation where skin temperature may be expected to change, the nonsensory contribution of such changes to the pain threshold must be distinguished from a true

analgesia or hyperalgesia. Measurements of the pain sensitivity of an individual requires about an hour, allowing a recommended minimum of 2 min. between stimuli and two determinations of the threshold, and skin temperature regulation.

(Manuscript received May 22, 1950)

References

1. Andrews, H. L. The effect of opiates on the pain threshold in post-addicts. *J. clin. Invest.*, 1943, **22**, 511–516.
2. Birren, J. E., Schapiro, H. B., and Miller, J. The effect of salicylate upon pain sensitivity. *J. Pharmacol. exp. Therap.*, 1950, **100**, 67–71.
3. Chapman, W. P., and Jones, C. M. Variations in cutaneous and visceral pain sensitivity in normal subjects. *J. clin. Invest.*, 1944, **23**, 81–91.
4. Chapman, W. P., Cohen, M. E., and Cobb, S. Measurements related to pain in neurocirculatory asthenia, anxiety neurosis, or effort syndrome: Levels of heat stimulus perceived as painful and producing wince and withdrawal reactions. *J. clin. Invest.*, 1946, **25**, 890–896.
5. Chapman, W. P., Finesinger, J. E., Jones, C. M., and Cobb, S. Measurements of pain sensitivity in patients with psychoneurosis. *Arch. Neurol. Psychiat., Chicago*, 1947, **57**, 321–331.
6. Hardy, J. D., Wolff, H. G., and Goodell, Helen. The pain threshold in man. *Amer. J. Psychiat.*, 1943, **99**, 744–751.
7. Hardy, J. D., Wolff, H. G., and Goodell, Helen. Studies on pain: Discrimination of differences in intensity of a pain stimulus as a basis of a scale of pain intensity. *J. clin. Invest.*, 1947, **26**, 1152–1158.
8. Kuhn, R. A., and Bromiley, R. B. Human pain thresholds determined by the radiant heat technique and the effect upon them of acetylsalicylic acid, morphine sulfate and sodium phenobarbital. *J. Pharmacal. exp. Therap.* (In press.)
9. Malmo, R. B., Shagass, C., Davis, J. F., Cleghorn, R. A., Graham, B. F., and Goodman, A. J. Standardized pain stimulation as controlled stress in physiological studies of psychoneurosis. *Science*, 1948, **108**, 509–511.
10. Miller, L. C. A critique of analgesic testing methods. *Ann. N.Y. Acad. Sci.*, 1948, **51**, 34–50.
11. Pfeiffer, C. C., Sonnenschein, R. R., Glassman, L., Jenny, E. H., and Bogoub, S. Experimental methods for studying analgesia. *Ann. N.Y. Acad. Sci.*, 1948, **51**, 21–33.
12. Sheard, C. Temperature of skin and thermal regulation of the body. In O. Glasser (Ed.), *Medical physics*. Chicago: Yearbook Publishers, 1944.
13. Wikler, A., Goodell, H., and Wolff, H. G. Studies on pain: The effects of analgesic agents on sensations other than pain. *J. Pharmacol. exp. Therap.*, 1945, **83**, 294–299.
14. Wolff, H. G., Hardy, J. D., and Goodell, H. Measurements of the effect on the pain threshold of acetylsalicylic acid, acetanilid, acetophenetidin, aminopyrine, ethyl alcohol, trichlorethylene, a barbiturate, quinine, ergotamine tartrate and caffeine: An analysis of their relation to the pain experience. *J. clin. Invest.*, 1941, **20**, 63–86.

A COMPARISON OF THE GALVANIC SKIN RESPONSES OF NORMALS AND PSYCHOTICS [1]

BY AUTAR S. PAINTAL

Department of Physiology, Lucknow University, India

Several investigations have been made on psychiatric patients by the electrodermal method. However, entire agreement between any two workers does not exist. For instance, in schizophrenic patients Peterson and Jung (6) found normal reactions, McCowan (3) greater, while Ödegaard (4) and Syz and Kinder (7) found diminished responses. Such differences in conclusions are in some measure due to variations in stimulus and technique used.

It was therefore decided to investigate further into the differences in the galvanic skin responses of normal individuals and psychotics, and to determine in the latter the degree, if any, of physiological impairment by a nociceptive stimulus such as faradic shock.

As shown by Haggard (1), a comparison of results is rendered difficult due to lack of a suitable unit of evaluating the individual responses. The present work also introduces a new method of evaluation which eliminates difficult formulae and conversion of the raw data into arbitrary derived measures.

Method

Circuit.—The basic circuit was the ordinary Wheatstone Bridge—the ratio arms consisting of two fixed resistances of 1,000 ohms and 10,000 ohms. A decade resistance box constituted the third arm, while in the unknown arm a calibrating unit was incorporated through one pole of a two-way key the other pole of which led off to the electrodes of the *S*.

The source of current was a 45-v. dry cell connected to the bridge through a fixed resistance. As each record was calibrated separately, thus avoiding incidental errors due to small variations in current strength, it was not considered necessary to maintain a constant current arrangement.

The B- and C-batteries of a General Electric electrocardiograph were removed, thus putting the amplifiers out of action, and the bridge was directly connected to the low period galvanometer. The deflections were optically recorded on sensitive paper moving at the rate of 1 in. per sec. Apart from these deflections (including those due to calibration), each record was automatically marked by vertical and horizontal lines spaced .04 in. from each other, the distance between two vertical lines representing .04 sec. The deflections of the galvanometer were quantitatively measured with respect to the horizontal lines. It should be noted that the calibrations were impressed on the same record.

Electrodes.—The zinc electrodes fixed in Lucite were of the type used by Haggard (2) and originally designed by M. A. Wenger. They were fixed securely by rubber bands to the palms of both hands at the commencement of each experiment. To maintain good contact, the electrodes were filled with electrode jelly impregnated with zinc sulphate.

Stimuli.—Two stimuli were used in each experiment: a faradic shock applied to the left leg, followed by a threat of shock.

Subjects.—The normal *S*s were selected at random from among medical students and from the normal attendants of patients visiting the outpatient department of the hospital. The psychotics were chosen at random from the inmates of the Agra Mental Hospital. They consisted chiefly of schizophrenic, manic-depressive, imbecile, and senile dementia patients.

Procedure.—*S* was connected to the bridge through the palms of both hands, and his electrical resistance was allowed to stabilize. After an average period of 10 min. an unexpected faradic shock was given on the left leg and the GSR recorded. About 2 min. later *S* was threatened with another shock and his response to the threat recorded. Nine hundred such

[1] This paper is part of a thesis presented to the Lucknow University, in partial fulfillment of the requirements for the M.D. degree. The writer is indebted to Professor Shyam N. Mathur for valuable advice throughout the investigation.

records, 450 normal and 450 of psychotics were obtained. Each record was calibrated separately, thereby avoiding all incidental errors due to variations in the strength of the current flowing through the bridge. The GSR's calculated in ohms were converted into micro log conductance (× 100) units.

Analysis

The data of normals and psychotics were analyzed separately but on similar lines.

Response to faradic shock.—All normal *Ss* were classified according to their initial electrical resistance of the skin, into resistance groups of 5,000 ohms each, ranging from 5,000 ohms to 45,000 ohms. The number of *Ss* registering resistances below and above the range mentioned were too few to be considered statistically. The mean of each group was determined, as was also the mean of its GSR's and its standard deviation.

As an additional measure, since a large majority of the psychotics examined were over 30 years of age, normals above 30 years of age were considered separately and as above.

The data of all psychotics were analyzed as a whole as in the case of normals. Where the number of observations was sufficiently large, the data of the clinical disease entities were analyzed separately. In the present paper, however, only the results obtained by a general analysis will be mentioned.

Response to threat.—Analyzed as for shock.

Ratio of GSR-threat to GSR-shock.—The ratios $\left(\frac{\text{GSR-threat}}{\text{GSR-shock}} \times 100\right)$ were determined both in ohms and log conductance units for each *S*. Also, the ratios of the mean GSR to shock and threat in each resistance group were determined.

Results

The results are shown graphically in Figs. 1 and 2. A linear relation between the initial resistance and the GSR to shock is apparent in Fig. 1. This is in contrast to the conclusions of Haggard, who found that the log conductance values tend to be constant around a particular level (1). The best-fitting straight line was therefore determined in each case both for shock and threat according to the following formulae:

$$\Sigma y - bn - A(\Sigma x) = 0$$

$$\Sigma xy - b(\Sigma x) - A(\Sigma x^2) = 0$$

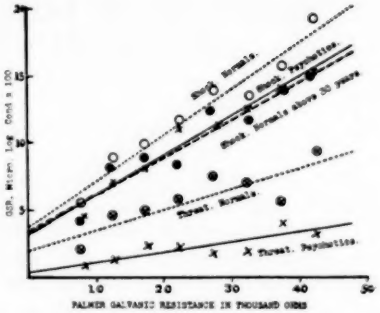

Fig. 1. Mean GSR's of normals and psychotics to faradic shock and threat

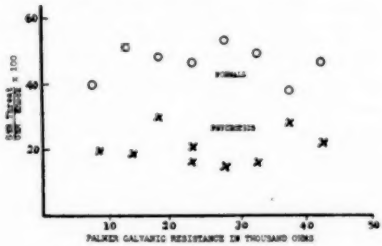

Fig. 2. Mean ratios of GSR's to threat and faradic shock of normals and psychotics at different levels of palmar galvanic resistance

from which are derived

$$A = \frac{(\Sigma x)(\Sigma y) - n(\Sigma xy)}{(\Sigma x)^2 - n(\Sigma x^2)}$$

and

$$b = \frac{(\Sigma x^2)(\Sigma y) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

where A and b occupy the respective positions in the following formula:

$$y = Ax + b\,.$$

The most striking feature is the close similarity of the responses to shock of normals and of psychotics, especially of normals above 30 years, in which case the best-fitting straight lines of the two practically coincide.

On the other hand, the response to threat is quite different. The level of reactivity is about half that for

shock. A greater variation in responses occurred, there being a larger number of negative GSR's, giving rise in some resistance groups to standard deviations greater than the corresponding mean GSR's. Further, in contrast to shock the gradient in psychotics is less than half that found in normals. The difference in the two stimuli is thus quite apparent, whereas one (shock) shows no difference in the responses of normals and psychotics, the other (threat) obviously does.

Threat-shock ratio.—The ratio $\frac{\text{GSR-threat}}{\text{GSR-shock}} \times 100$ further brings out the difference in the response to threat. In Fig. 2, two levels of reactivity are shown. Apart from a few deviations, possibly due to the small number of observations in these groups, the points tend to settle around two levels—50 per cent and 20 per cent for the normal and psychotic groups, respectively. The highest ratio in psychotics is considerably lower than the lowest in normals. It may be concluded that the average response to threat is about 50 per cent of that to shock in normals and about 20 per cent in psychotics, i.e., the response to threat in normals relative to shock is 2.5 times that in psychotics.

Several normal *Ss* were seen who gave diminished or negative responses to threat. These *Ss* also gave low or negative responses to shock. Similarly, several psychotics gave responses to threat much larger than the mean normal responses. Their response to shock was correspondingly large.

Discussion

The results indicate that the physiological mechanism responsible for giving rise to the galvanic skin response is not impaired in psychotics, as is shown by the normal responses to shock. The response so obtained may be considered to be one of several reflex sympathetic reactions to a nociceptive stimulus. The value of shock or some such stimulus would thus seem to lie in determining the integrity of the various autonomic paths giving rise to the GSR since threat—though in itself a powerful stimulus—or any other psychological stimulus very often gives negative responses. In the second place, the results indicate that in psychotics, as expected, there is a low level of appreciation of threat, probably due to an impairment of judgment. Finally, the importance of the nature of the stimulus used in experimentation cannot be overestimated.

Diminished, negative, or unduly large responses which occur where least expected are a common occurrence in any series of GSR's. In the present work a diminished or negative response to threat would probably indicate a lack of appreciation of the threat. On the other hand, an unduly large response in psychotics would probably preclude the occurrence of such a deficiency. It is, therefore, clear that conclusions from responses as such might be misleading.

The sympathetic response, quantitatively, like any other physiological response, differs in different individuals, and therefore galvanic skin responses of the same magnitude at the same level of palmar galvanic resistance in different people could not have the same significance. Consequently, it is suggested that the response to any particular stimulus should not be judged with reference to a mean obtained in a set of controls but with reference to the highest level of sympathetic activity that can be called forth in that individual. Therefore, if by some method the

maximum galvanic skin response of an individual be determined (by faradic shock in the present work), a base might then be obtained from which all other responses to different stimuli could be judged. The ratio, extensively used in the present work, is an attempt in this direction.

An example will make this clear. The data of case No. 351 (**5**), a schizophrenic, are as follows:

Initial resistance = 16,880 ohms
GSR-shock = 6,210 ohms: 19.93 log cond.
GSR-threat = 1,680 ohms: 5.10 log cond.

$$\frac{\text{GSR-threat}}{\text{GSR-shock}} \times 100 = \begin{cases} 27.65 \text{ per cent (calculated in ohms)} \\ 25.59 \text{ per cent (calculated in log cond.)} \end{cases}$$

The normal response to threat at the initial resistance level of 16,880 ohms, as determined from Fig. 1, would be approximately 4.63 log conductance units. If the GSR to threat (5.10 log cond.) of the case under consideration be judged with respect to the normal response of 4.63 log conductance units, it would be concluded that the response of the patient was above normal and that probably the patient appreciated the significance of the threat. But on applying the threat-shock ratio, it becomes evident that the response is far below the level of normal *Ss* (about 50 per cent), being 25.59 per cent.

Similarly, there are a large number of diminished responses which would be interpreted as abnormal were the ratio not applied to them.

Another advantage of the ratio method is that conversion of the ohms scale into derived units (**1**) can be dispensed with inasmuch as the general level of electrical resistance from which GSR's are obtained in *S* is usually the same. This has been demonstrated by the close similarity of the ratios obtained in ohms and in log conductance units (**5**).

Summary

1. The galvanic skin responses of 450 normals and 450 psychotics to faradic shock and threat of shock have been compared.

2. The results show that the physiological mechanism responsible for giving rise to the GSR to shock is not impaired in psychotics. However, there was markedly less response to threat of shock in psychotics than in normal *Ss*.

3. It is suggested that the value of a GSR should be judged with reference to the maximum response obtained in an individual, since in itself the GSR is often misleading. Further, by expressing a GSR as a percentage of the maximum response, conversion of the raw data into derived measures can be dispensed with.

(Manuscript received May 30, 1950)

References

1. Haggard, E. A. Experimental studies in affective processes: II. On the quantification and evaluation of "measured" changes in skin resistance. *J. exp. Psychol.*, 1945, **35**, 46–56.
2. Haggard, E. A., and Gerbrands, R. An apparatus for measuring continuous changes in palmar skin resistance. *J. exp. Psychol.*, 1947, **37**, 92–98.
3. Landis, E., and DeWick, H. N. The electrical phenomenon of the skin (psychogalvanic reflex). *Psychol. Bull.*, 1929, **26**, 64–102.
4. Ödegaard, Ö. The psychogalvanic reactivity in affective disorders. *Brit. J. med. Psychol.*, 1932, **72**, 132–150.
5. Paintal, A. S. The electrical resistance of the skin in normal and mentally diseased individuals. Unpublished M.D. thesis, Lucknow Univ., 1949.
6. Peterson, F., and Jung, C. G. Psychophysical investigations with the galvanometer and plethysmograph in normal and insane individuals. *Brain*, 1897, **30**, 153–218.
7. Syz, H. C., and Kinder, Elaine F. The galvanic skin reflex: Further aspects in psychopathological groups. *Arch. Neurol. Psychiat.*, 1931, **26**, 146–155.

STUDIES OF THE EFFECT OF CHANGE OF DRIVE: I. FROM HUNGER TO THIRST IN A T-MAZE [1]

BY HOWARD H. KENDLER AND SEYMOUR LEVINE

New York University

In 1946 Kendler (5) reported experimental evidence which indicated the marked influence a change of drive had upon the performance of albino rats in a T-maze. Thirty-two animals had received three free-choice trials daily in a T-maze having one end-box baited with dry food. After the *S*s reached the criterion of four successive days of correct choices, their motivation was shifted to thirst. On the first trial under thirst motivation 78 per cent of the *S*s chose the side opposite to the one they had previously taken for 12 successive trials. This phenomenon of choosing the side opposite to the one which the animals had been consistently choosing was labeled "switching."

The explanation initially offered (5) was that the switching behavior was similar to regression. By reasoning within an S-R reinforcement theoretical framework, it was assumed, first, that when the hunger drive stimulus was eliminated, little, if any, habit strength was generalized to the *new drive stimulus component* (thirst) of the choice-point stimulus compound. Secondly, it was assumed that the new thirst drive stimulus acted in a manner similar to an "external" inhibitor. These two factors would have as their consequence the "wiping-out" of the effects of training to the food side. Since the animals had been trained against their position preference (i.e., food had been placed on the side of the maze that *S* least preferred during a series of position preference trials) it would be expected that the *S*s would regress to their initial position preference when their motivation was shifted to thirst.

An experimental test of this explanation failed to support the analysis. Kendler, in an unpublished study, trained two groups of albino rats to choose the side leading to food in a T-maze. One group found food in the end-box on the side consistent with their position preference; the other group had their food in the end-box on the side opposite to their position preference. Approximately half of the *S*s in both groups switched when the motivation of the *S*s was shifted to thirst. This suggested that the *S*'s initial position preference did not influence the amount of switching.

The purpose of this paper is to report the results of an experiment which was designed to test another possible explanation of the switching phenomenon. This hypothesis was suggested by some data reported in a study (7) which tested certain implications of Tolman's sign-gestalt theory of trial-and-error learning. In this study animals were trained in a T-maze having food in one end-box and water in the other. In order to equate the experience of the animals

[1] This paper reports four duplicated studies. The reason for the duplication was that the initial studies were teasingly suggestive but not statistically significant. The first two studies were conducted at the University of Colorado and were aided by a grant-in-aid from the Council on Research and Creative Work of the Graduate School. The second two studies, which were part of the junior author's M.A. thesis, were conducted at New York University. The authors are indebted to Mrs. Phyliss A. Alderman, who collected the data of the second study.

with the contents of each goal box, a system involving two free and two forced trials was used daily. On the last day of training (seventh day) the *thirsty* animals averaged 4.8 sec. on their free-choice trials (all to water) and 78.6 sec. on their forced trials to food. The *hungry* animals had a mean running time of 6.3 sec. on their free-choice trials (all to food) as compared with a mean running time of 9.9 sec. on their forced trials (all to water). It appears, therefore, that thirsty animals display a stronger aversion to dry food than hungry animals do to water.

This assumption of a strong aversion toward dry food when animals are thirsty is the cornerstone of our proposed hypothesis. On the descriptive level, it is assumed that when *S*s are made thirsty, following their training to go to dry food, they switch their spatial response in order to avoid the dry food. The basic theoretical mechanism underlying the foresightful behavior is assumed to be the anticipatory goal response.

It is reasoned that when *S*s were hungry and food was in one of the end-boxes, the cues emanating from the food-box, and the alley leading to the food, became associated with anticipatory eating responses (masticatory movements, salivation, etc.). The proprioceptive stimuli resulting from these anticipatory goal responses became themselves, in turn, conditioned to entering and continuing locomotion in the alley leading to the food-box. Initially, only the hunger drive stimulus and the food would evoke the anticipatory eating responses. As learning to go to food progressed, however, these anticipatory responses became conditioned to certain features of the goal box containing food and then "worked backward" until cues at the entrance to the correct alley were capable of evoking some component of the anticipatory eating response. When the animals were made thirsty, the cues of the maze elicited these anticipatory eating responses. It is assumed that the anticipatory eating responses of dry food are, in some manner, incompatible with the resultant condition of water deprivation (probably conflict with anticipatory drinking responses). It would be expected that this "incompatibility" at the choice point would be reduced by the *S*s going in the direction *away* from the end-box baited with food. The cues in the incorrect path of the maze had not been directly conditioned to the anticipatory eating responses, and, consequently, the greater the penetration into this path, the greater would be the relief from the "conflict state" postulated to exist.

The present group of studies was designed to test one implication of the above analysis. It would be expected that any factor which would increase the intensity of the anticipatory eating response at the choice point would increase the probability of switching behavior under thirst motivation. One such factor would be the number of reinforcements of the response to food. With the anticipatory eating habit getting more numerous reinforcements, it would be stronger and would have more opportunity to work backwards to the choice point. This would lead to the prediction that the more an animal is trained to go to one side of a T-maze to get food, the more likely it would be (up to a certain point) to choose the other side when made thirsty.

Method

Experimental design.—The experimental design for all four studies was similar. A comparison was made between the amount of switching exhibited by *S*s who had different

strengths of turning habits to food. The strengths of the turning habit were manipulated by using different criteria of learning to go to the food side. Group 1 was required to reach a criterion of one day of correct choices to food, while Group 4 was required to reach a criterion of four successive days of correct choices. On the day following the achievement of the criterion of learning, all *S*s were made thirsty and subjected to the T-maze situation.

Subjects.—The *S*s for these investigations were naive male albino rats whose ages at the beginning of the experiment were approximately 85 days. The 18 rats of the first study were drawn from the colony maintained by the Chemistry Department, while the 33 rats of the second study were from the Department of Psychology's strain, of the University of Colorado. The 85 rats of the last two studies were purchased from the Albino Farm, Red Banks, New Jersey.

Apparatus.—A single-choice T-maze was used. A description of a similar maze is given in a previous paper (6). For all studies except the second the entire maze was painted a flat black. During the second study the left half of the maze was painted black, and the right half was painted white.

Preliminary training.—In the first study after each *S* had been handled from 3 to 5 min. daily for approximately a week, a preliminary training period in a straightaway was begun. Following this the *S*s were placed individually in the T-maze. Their initial choices were recorded. For the next three days the *S*s received three daily free-choice, non-rewarded trials. The position preference of the *S*s of the first study could be determined from their performance on these ten nonrewarded trials.[2]

The preliminary training in the second study did not include the straightaway experience. The *S*s were placed individually in the T-maze for 30 min. Their initial choice was recorded. On the following day they received two free-choice, nonrewarded trials. These *S*s, therefore, had a total of three trials which could be used as a measure of position preference.

The preliminary training of the third and fourth study included experience in a straightaway. The *S*s, following this training, were subjected to three free-choice, nonrewarded trials.

Training series.—During the training series *S*s were motivated by hunger. Food was present in one end-box, while the other end-box remained empty. In the initial two studies 10 pellets of Purina Dog Chow were placed in the baited end-box, whereas in the latter two studies this number was increased to 15.

The hunger motivation was controlled by having approximately 10 gm. of Purina Dog Chow available to each *S* for 1 hr. daily. This feeding period terminated approximately 22 hr. prior to the daily experimental trials. Water was constantly available in the home cages throughout the training series.

The training series consisted of three free-choice trials daily until the criterion of learning was achieved. When it chose the correct side, *S* was permitted to eat for 20 sec.; when it chose the incorrect side, it was confined to the empty end-box for 30 sec. Food was placed behind the empty end-box to control for any possible olfactory cues.

In all studies except the first, the *S*s were randomly divided into two groups. The *S*s in Group 1 were trained until a criterion of one day of correct choices (at least three successive errorless trials) was achieved. The *S*s in Group 4 were trained until they reached a criterion of four days of correct choices (at least 12 successive errorless trials).

Test trial.—The *S*s in both groups, during the test trials, were thirsty while being satiated for food. This motivational condition was manipulated by giving the animals approximately half of their daily rations of Purina Dog Chow about 20 min. after their last training trial. They had access to water for 1 hr. Following the end of the hour the water was removed and *S*s were presented the remainder of their daily ration. One hour prior to the first test trial a dish of Kibble meal was placed in the home cages.

Results

As would be expected, the *S*s achieving the criterion of learning of four successive days of correct responses required significantly more reinforcements to reach their criterion of learning than did those requiring one perfect day. The mean number of reinforcements (excluding the criterion trials) required by the one-day groups ranged from 2.3 to 3.0, and the

[2] The first study was instigated during the terminal part of the investigation aimed at discovering whether the position preference of white rats influenced switching behavior. Inasmuch as the data strongly suggested that position preference was not a relevant variable, it was decided to test the "avoidance hypothesis" by subjecting the last five *S*s to the procedure involving a one-day criterion of learning.

TABLE I

COMPARISON OF THE NUMBER AND PER CENT OF *Ss* IN GROUPS 1 AND 4 SWITCHING SPATIAL RESPONSE UNDER THIRST MOTIVATION

Study	Switching Response								
	Group 1			Group 4			Difference		
	N	Number	%	*N*	Number	%	σ_{diff}	*CR*	*P**
I	5	1	20.00	13	7	53.85	22.61	1.50	.07
II	16	4	25.00	17	7	41.18	16.11	1.00	.16
III	20	7	35.00	21	10	47.62	15.25	0.83	.20
IV	22	7	31.82	22	11	50.00	14.57	1.25	.11
Total	63	19	30.16	73	35	47.95	8.22	2.16	.015

* Single-tailed test.

range of the four-day groups was between 5.1 and 13.2 reinforcements.

Table I presents the major results. The data reveal the influence that the strength of the habit acquired under hunger had upon the amount of "switching." Only one tail of the sampling distribution of differences in percentage is considered. This was done because the hypothesis tested assumed a positive relationship between the amount of training and the amount of switching. In none of the four studies was the obtained difference significant beyond the .05 level. In all studies, however, those animals that achieved a higher criterion of learning had a greater amount of switching.

It is interesting to note the striking consistency in the results of the various independently conducted experiments. Since all studies utilized essentially the same experimental procedure, it appears justified and advantageous to combine the data. One method by which the data can be combined is to treat the percentage of switching of each group for each study as separate scores. Then a comparison, by means of a *U*-test (**12**), may be made between the four scores of Group 1 and the four scores of Group 4.[3] The specific procedure used in making this comparison is to rank the scores (percentages) of the two distributions and determine how many times a score in the lower distribution (in this case, Group 1) is higher than scores in the upper distribution. This procedure yields the value of *U*. Since there is no overlap between the two distributions, *U* is equal to zero. This value, when treated in the appropriate formula (**12**), is significant at the .014 level. The results of this statistical analysis appear, therefore, to be consistent with our theoretical expectations.

Another set of data which are relevant to the theoretical analysis is the relative disturbances of the running times instigated by the shift of the drive from hunger to thirst. Table II compares the running time on the first trial under thirst drive with that of the first trial of the preceding day. For Group 1 the first trial of the preceding day was the first criterion trial, while for Group 4 this trial was the tenth criterion trial. A statistical technique to treat data from skewed distributions was used (**1**). It

[3] The authors are indebted to Dr. Toby Oxtoby, who suggested, and taught them, this test.

should be noted that there was a significant increase in running time for Group 4 in all studies. This was not true for Group 1. In fact, in two of the studies there was a decrease in running time. These data suggest, therefore, that the running time of the groups reaching the high criterion of learning was more disturbed than that of the groups reaching the low criterion. Such findings are consistent with our theoretical expectations.

Discussion

The basic design of the four studies consisted in training animals to acquire a turning habit in a T-maze situation. Following the achievement of a criterion of learning, the motivation under which the habit was acquired was eliminated and another substituted.

One would expect, using the postulates of Hull's behavior system (4), that the tendency of *Ss* to persist in the execution of the learned response under the new drive condition would be somewhat reduced. By changing the drive, the stimulus compound which had been associated with the appropriate response would be modified. The drive stimulus[4] component, which Hull assumes, for the sake of convenience, acquires 50 per cent of the habit loading, would be eliminated. But because of certain constant environmental stimulus objects, some habit strength would be generalized to the stimulus compound involving the new drive stimulus. Although one would not expect the *Ss* to persist in the 100 per cent level of execution of the learned turning habit under the new drive condition, it would be expected that they should respond above the chance level. The results of the *Ss* in Group 4 that persisted at only a chance level in the execution of the acquired turning response would appear, at first glance, to be inconsistent with certain implications of Hull's theory.

The possibility remains, however, that Hull's analysis is valid; the shift from hunger to thirst introduces additional factors which conceal the basic relationship between maze performance and change of drive.

It was assumed that when hungry animals are trained to choose the side of a T-maze leading to dry food, certain maze cues become conditioned to the anticipatory eating response of dry food. It was further assumed that when the drive state is shifted to thirst, the animals tend to avoid the cues which most strongly elicit the anticipatory eating response. The proposed explanation led to the expectation that those factors which

TABLE II

Comparison of Mean Running Time on Trial 1 under Thirst Motivation with Trial 1 of the Preceding Day

Study	Mean Running Time (Sec.)					
	Group 1			Group 4		
	Trial 1 of Prec. Day	Trial 1 on Thirst	*P*	Trial 1 of Prec. Day	Trial 1 on Thirst	*P*
I	5.4	10.8	.05	2.9	9.3	.01
II	9.8	8.4	—	4.3	15.4	.01
III	14.4	7.8	.05	4.4	12.2	.01
IV	16.5	20.9	—	8.9	41.9	.01

[4] Drive stimulus (S_D) is an intervening variable defined as some function of the feeding schedule which is under the control of the experimenter. This concept implies that the drive itself acts as a stimulus and therefore can become associated with a response. It should be clearly distinguished from drive (D), which is an intervening variable defined as some function of all specific needs of the moment. It acts in a multiplicative manner with habits to determine behavior.

would increase the intensity of the anticipatory eating response at the choice-point situation would also increase the probability of "switching" behavior under thirst motivation. This expectation was confirmed by the results of the present studies. The animals that reached a relatively low criterion of learning (thereby having weaker anticipatory eating responses) revealed less switching behavior.[5]

The results of these studies have certain implications for a recent controversy (**11, 16**) as to whether or not animals are capable of learning the location of unwanted goal objects. Leeper (**11**), in his criticism of the initial Iowa latent learning studies (**7, 8, 17**), mentions some experimental results which he believes to be inconsistent with the Iowa studies. In his experiment food was in one goal box and water was in the other. He writes: "The rats were given four or five days of training under hunger at the start of their work . . . before their first shift to thirst. On the first trial after that shift, 13 of the 23 rats went to the now correct side." Leeper interpreted such results to support his contention that when his animals went to the water side, when hungry, they formed a cognition of the location of water. When made thirsty, they would utilize that cognition by choosing the side leading to the water. However, according to the findings of the present study, similar results might have been obtained even if *no water* were present during the initial four or five days of training, i.e., these 13 animals did not switch because they were utilizing when thirsty a cognition of the location of water; they switched in order to reduce the conflict state induced by anticipatory goal responses of eating dry food when thirsty.

In a similar manner, these conclusions may be invoked to explain some "intriguing findings" recently reported by Meehl and MacCorquodale (**13**). They trained animals in a T-maze which contained food in one end-box and water in the other. Following a rather complicated experimental procedure, these experimenters found "that the rats who had been reinforced with water while thirsty, tended, when they are subsequently made hungry, to continue going to the reinforced side." Such a response was designated as a "Hullwise response" inasmuch as it was the response to that side of the maze which had been reinforced. They found, however, that rats reinforced with food tended when subsequently made thirsty to choose water, i.e., behave "Tolmanwise." As in Leeper's experiment, the behavior of "Tolmanwise" animals may not of necessity indicate that a cognition of the location of water had been formed when the animals were hungry. Such behavior may be similar to the switching response reported in this paper. The thirsty animals that behaved "Tolmanwise" chose the water side not because they "knew" the water was there, but, rather, because that response reduced the conflict state postulated to exist between anticipatory eating responses of dry food and some resultant condition of thirst.

[5] The proposed explanation may be tested in many ways. A few of the deductions would be that less switching should occur 1) when the reward during the initial training is a wet mash rather than dry food, 2) when the motivation is shifted to a weak thirst drive, and 3) when the motivation is shifted from thirst to hunger rather than from hunger to thirst. It would also be expected that animals would learn more quickly to choose the shorter of two paths to water when the longer path requires the animals to run over dry food. This would follow because the distance discrimination would be facilitated by the development of an avoidance response to the longer path. The results of some unreported "pilot studies" appear to support the first three deductions, and Walker's findings (**18**) tend to be consistent with the last expectation.

The findings of the present study may appear at first glance to be inconsistent with those reported by Spence and Lippitt (17) and by Kendler (7). These investigators trained animals under one uniform motivational condition (either hunger or thirst) in a single-choice maze situation in which water was present in one end-box while food was in the other. Following the completion of training during which time the animals had familiarized themselves with the contents of both goal boxes, their motivation was shifted in order to ascertain whether they had learned the location of the goal object for which they were not motivated during the training trials. In both investigations, *all* animals under the new motivational state persisted in the execution of the habit which had led to the end-box containing the commodity for which the animals had been motivated during the training trials. Why did the animals in these studies fail to show any "switching" behavior? The answer, it is believed, lies in the methodology used by Spence and Lippitt, and by Kendler, to equate the experience of the animals with the contents of both end-boxes. When animals were hungry, they naturally tended to choose with a successively higher degree of accuracy, the side which led to the food. As training progressed, the animals evidenced more and more aversion to being forced down the alley which led to the unwanted goal object and *away* from the desired commodity. It is quite possible that a strong avoidance response to that side was developed and persisted when the animals' motivation was shifted. If such were the case, the animals, during the test trials, may not have been choosing the side on which they had been rewarded, but, rather, they may have been avoiding the side on which they had been frustrated. Such an interpretation, although differing from that given by those authors, still does not invalidate their conclusions. The fact remains that the animals in these studies had some of their receptors stimulated by the unwanted goal object and, later, when motivated for that commodity, did not respond appropriately. Other recent experiments (2, 9), in which the methodology precluded the development of avoidance responses, provided similar results, i.e., animals were incapable of learning the location of "unwanted" food when motivated by a thirst drive based upon 20 to 22 hr. of water deprivation.

The data of these studies and their accompanying interpretation possess certain general implications for the field of learning. First, it should be noted that the results of these studies definitely reveal that stimuli arising from drive conditions (or those that arouse the drive condition) become an important component of a habit structure. It therefore becomes extremely important in describing any habitual mode of response of an organism to specify the drive condition under which the habit was acquired. Secondly, it becomes necessary for those experimenters who utilize a change-of-drive technique in their experimental procedure to provide adequate controls to evaluate the influence of their experimental variables. Thirdly, it should be noted that the "switching" response is essentially a novel response, i.e., the response "emerges" following a certain type of training and a change in the motivational condition of the animal. An attempt has been made to explain this novel behavior in terms of an S-R theory. One of the major criticisms leveled at such theories is that they are incapable of explaining such novel behavior. The work of Hull (3),

Miller (**14**), and Spence (**15**), as well as the interpretation of the present findings, reveals the weakness of such a criticism.

Summary

Kendler (**5**) has reported evidence indicating that following the acquisition of a turning habit under hunger motivation and food reward in a T-maze situation, the invocation of a thirst drive would cause the animals to "switch" their spatial response, i.e., turn left, if the turning habit acquired under hunger was to the right. The purpose of the present study was to investigate the influence of the amount of training of the turning habit acquired under hunger motivation and food reward. This was accomplished by having groups of rats reach a criterion of either one or four days of correct turning responses. The results indicate that the greater the amount of training on the initial turning habit, the greater the amount of "switching" behavior exhibited under the new thirst drive.

The results are interpreted to be consistent with the theory which instigated the present research. It assumes that there is something incompatible between fractional anticipatory eating responses of dry food and some resultant condition of water deprivation. The animals "switch" their turning response in order to achieve some relief from this conflict situation.

(Manuscript received June 6, 1950)

References

1. Festinger, L. A statistical test for means of samples from skewed populations. *Psychometrika*, 1943, **8**, 205–210.
2. Grice, G. R. An experimental test of the expectation theory of learning. *J. comp. physiol. Psychol.*, 1948, **41**, 137–143.
3. Hull, C. L. Mind, mechanisms and adaptive behavior. *Psychol. Rev.*, 1937, **44**, 1–32.
4. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
5. Kendler, H. H. The effect of change of drive upon T-maze performance. *Amer. Psychologist*, 1946, **1**, 246 (Abstract).
6. Kendler, H. H. The influence of simultaneous hunger and thirst drives upon the learning of two opposed spatial responses of the white rat. *J. exp. Psychol.*, 1946, **36**, 212–220.
7. Kendler, H. H. An investigation of latent learning in a T-maze. *J. comp. physiol. Psychol.*, 1947, **40**, 265–270.
8. Kendler, H. H. A comparison of learning under motivated and satiated conditions in the white rat. *J. exp. Psychol.*, 1947, **37**, 545–549.
9. Kendler, H. H., and Mencher, Helen C. The ability of rats to learn the location of food when motivated by thirst—an experimental reply to Leeper. *J. exp. Psychol.*, 1948, **38**, 82–88.
10. Leeper, R. The role of motivation in learning: A study of the phenomenon of differential motivational control of the utilization of habits. *J. genet. Psychol.*, 1935, **46**, 3–40.
11. Leeper, R. The experiments by Spence and Lippitt and by Kendler on the sign-gestalt theory of learning. *J. exp. Psychol.*, 1948, **38**, 102–106.
12. Mann, H. B., and Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Stat.*, 1947, **18**, 50–60.
13. Meehl, P. E., and MacCorquodale, K. A further study of latent learning in the T-maze. *J. comp. physiol. Psychol.*, 1948, **41**, 372–396.
14. Miller, N. E. A reply to a 'Sign-gestalt or conditioned reflex?'. *Psychol. Rev.*, 1935, **42**, 280–292.
15. Spence, K. W. The differential responses in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, **44**, 430–444.
16. Spence, K. W., and Kendler, H. H. The speculations of Leeper with respect to the Iowa tests of the sign-gestalt theory of learning. *J. exp. Psychol.*, 1948, **38**, 106–109
17. Spence, K. W., and Lippitt, R. An experimental test of the sign-gestalt theory of trial and error learning. *J. exp. Psychol.*, 1946, **36**, 491–502.
18. Walker, E. L. Drive specificity and learning. *J. exp. Psychol.*, 1948, **38**, 39–50.

DELAY OF REWARD AND HUMAN VERBAL LEARNING

BY IRVING J. SALTZMAN

Indiana University

A topic which has received much attention from learning theorists is delay of reward and its effect on learning. A large number of rigorously designed and well-controlled experiments on this topic has appeared in the literature (4). The efforts of the investigators in this area have thus far been almost exclusively directed to the learning of subhuman animals—primarily laboratory rats. Yet, the thoroughness of the animal experimentation makes it tempting to extrapolate to human behavior. Before any such generalizations or extrapolations are completely justified, however, some experimentation on the topic should be carried out with human subjects. This study represents two attempts to investigate the effects of delay of reward on human learning.

In attempting to study this problem at the human level we are immediately faced with two problems. The first is deciding on the kind of learning material to employ; and the second is deciding on an adequate reward for human beings—a reward that can be conveniently delayed. As a beginning, it was felt that a verbal maze, presented by means of a memory drum, could serve adequately as the learning task. Reward in the situation could be achieved by presenting information to the *S*s concerning the correctness of their responses. Following each response *S* could be informed whether or not the response had been correct. Informing *S* that a correct response had been made would be the reward.

Experiment I

Subjects.—Forty-four Indiana University freshman and sophomore men and women were randomly divided into two groups of 22. None of the *S*s had ever served previously in a psychological experiment.

Apparatus and method.—The learning task was a verbal maze consisting of 10 pairs of four-place numbers. One group of *S*s learned the maze with immediate reward, and the other, with delayed reward. One four-place number in each pair was previously selected by *E* as the "correct" number. The maze was presented to the *S*s individually by means of an electrically controlled memory drum. Each pair of numbers appeared in an aperture in the drum for 6 sec. During the 6-sec. period *S* was required to state which number was correct by saying "left" to indicate the number on the left, or "right" to indicate the number on the right. For the group with immediate reward, the correct answer appeared immediately after the 6-sec. presentation period. If the left-hand number was correct, the letter *L* appeared in the drum aperture; if the right-hand number was correct, the letter *R* appeared. The letters *L* and *R* remained in the aperture for 6 sec. and then were replaced by the next pair of numbers, etc. For the group with delay of reward, a blank period of 6 sec. intervened between the removal of the numbers and the appearance of the correct answer. The immediate reward group, then, responded approximately every 12 sec. (6-sec. presentation + 6-sec. answer) and the delayed reward group responded roughly every 18 sec. (6-sec. presentation + 6-sec. delay + 6-sec. answer).

Each complete presentation of the list was followed by a 30-sec. rest period. The pairs of numbers always appeared in the same left-right position and in the same order on successive trials. Learning trials were continued until each *S* made two consecutive perfect performances of the complete list. The *S*s were required to respond on the very first trial, although they had no sound basis for making their choices. This was done in order to familiarize *S*s with the procedure—so that they would learn to make their responses within the 6-sec. presentation period.

Before the experiment was begun, *S*s were told that the purpose of the experiment was to determine the speed with which they could

learn the numbers, and they were asked to try to learn the numbers as rapidly as possible. Each response made during the learning trials was recorded by *E*. The errors and trials required to meet the criterion were determined. Each *S* was questioned following the experiment to discover the manner in which he had approached the task.

Results and discussion.—The group that learned the maze with immediate reward required a mean of 6.68 trials to satisfy the criterion, and the group that learned with a 6-sec. delay of reward required a mean of 5.95 trials. This difference is not statistically significant ($t = .94$). The mean total number of errors was 6.90 for the immediate reward group and 6.05 for the delay group, exclusive of errors made on the first trial. This difference is also not statistically significant ($t = .49$).

At first glance it appears that delaying the reward with human *Ss* has no effect on their learning. But an analysis of *Ss*' comments indicates that such a conclusion is probably not warranted. Of the 44 *Ss*, 37 reported that they were not bothering with the numbers themselves, but were actually learning only a pattern of correct responses. Most of the *Ss* could recall only two or three of the correct numbers following the two perfect criterion trials, but they could readily recite the pattern of correct responses, such as left-right-right-left. In other words, the design of the experiment and *Ss*' approach to the problem were such that *E* was not actually controlling the time between the stimulus-response occurrence and the reward. The major difference between the two groups of *Ss* was that the letters *L* and *R* were being presented in more rapid succession in one case than in the other. In order to eliminate pattern learning and carry out a study where a more precise control of delay of reward was attained, a second attempt (Experiment II) was made.

Although the results of this experiment are of little significance with regard to the problem of delay of reward and human learning, they are of interest, nevertheless, in a different regard. Hovland (3) recently reported the results of an experiment showing that the rate of learning nonsense syllables depends on the rate at which they are presented to the *Ss*. He found that when the syllables were presented so that the *Ss* responded every 4 sec., learning was significantly more rapid than when the *Ss* responded every 2 sec. These results are in agreement with earlier studies in this area (1, 2, 5). In the present study, it was found that *Ss* who responded every 18 sec. (delayed reward group) learned in fewer trials than *Ss* who responded every 12 sec. (immediate reward group). The difference between the groups, however, is not statistically significant. We may tentatively conclude that decreasing the rate of presentation beyond rates of about 15 or 20 sec. per item will not result in more rapid learning, with learning of the type involved in this experiment.

Experiment II

Subjects.—Again, 44 Indiana University freshman and sophomore men and women were randomly divided into two equal groups, a delay-of-reward group and an immediate reward group. None of the *Ss* had served previously in psychological experiments.

Apparatus and method.—The memory drum was used, and essentially the same procedure was followed in this experiment as was employed in the first experiment, with the following exceptions: (*a*) The numbers appeared in a different order on successive trials. Four different orders were used. (*b*) The left-right position of the numbers within a given pair was not always the same. These changes eliminated the practicality of learning a pattern of left-right responses, since a number which happened to be on the right on one trial, would not necessarily be on the right on succeeding trials. (*c*) Only six pairs of four-place numbers were used, instead of ten.

Results and discussion.—The group learning with immediate reward required a mean of 9.95 trials and made a mean of 14.50 total errors; the group learning with the 6-sec. delay required 14.09 trials and made 21.95 total errors. Each difference between groups is significant at the .01 level of confidence (t = 2.84 and 2.92, respectively).

On the basis of the results of this second experiment we may conclude that when the reward is effectively delayed with human *Ss* (where the reward is defined as informing the *S* that a correct response has been made), learning will be slower than in instances where the reward is not delayed. Although this conclusion may be drawn from the results of the present study, it is felt that a statement of the complete relationship between human learning rate and temporal delay of reward cannot be made until further experimentation in this area is completed. In the first place, it should be noted that not only was the *reward* delayed in this experiment, but information about incorrect responses was delayed also. The outcome of this experiment, then, may be partly due to delayed information about incorrect responses, and not completely due to delayed reward. Secondly, in carrying out the experiment, *E* noticed that almost invariably the *Ss* repeated or rehearsed the selected number of a pair. Since the delayed reward group was able to rehearse longer than the immediate reward group, it appears that rehearsing in advance of knowledge of the correctness of the response may lead to slower learning.

Although the present experiment is limited to a verbal learning task of moderate difficulty, the results suggest that an important variable affecting human learning as well as subhuman learning is temporal delay of reward.

Summary and Conclusions

Two attempts were made to determine the effect of delaying reward on the speed of human learning. The learning task was a verbal maze, and information that a response was correct was the reward. Forty-four college students served as *Ss* in each experiment. The first experiment, although it provided no information concerning the effects of delayed rewards on learning, did provide information with regard to the speed of presenting learning material and the rate of learning. Reducing the presentation time from 18 sec. between items to 12 sec. does not affect the learning rate significantly, although the direction of the effect is consistent with other studies. The second experiment showed that when reward is effectively delayed, learning is slower than when the reward is not delayed.

(Manuscript received June 21, 1950)

References

1. Hovland, C. I. Experimental studies in rote-learning theory. II. Reminiscence with varying speeds of syllable presentation. *J. exp. Psychol.*, 1938, **22**, 338–353.
2. Hovland, C. I. Experimental studies in rote-learning theory. III. Distribution of practice with varying speeds of syllable presentation. *J. exp. Psychol.*, 1938, **23**, 172–190.
3. Hovland, C. I. Experimental studies in rote-learning theory. VIII. Distributed practice of paired associates with varying rates of presentation. *J. exp. Psychol.*, 1949, **39**, 714–718.
4. Melton, A. W. Learning. *Ann. Rev. Psychol.*, 1949, **1**, 9–30.
5. Melton, A. W., and Stone, G. R. The retention of serial lists of adjectives over short time-intervals with varying rates of presentation. *J. exp. Psychol.*, 1942, **30**, 295–310.

THE RATE OF LEARNING A TONE–NO-TONE DISCRIMINATION AS A FUNCTION OF THE TONE DURATION AT THE TIME OF THE CHOICEPOINT RESPONSE [1]

BY M. U. ENINGER

Carnegie Institute of Technology

The present study is concerned with one of the variables determining the capacity of a stimulus to acquire associative connections with a response. According to Hull's (1) principle of the dynamic state of stimuli, those stimuli which have continued to act upon receptor mechanisms for long durations have a diminished capacity for acquiring habit loadings. The principle implies the theoretical proposition that the magnitude of an increment of habit strength resulting from a stimulus-response reinforcement sequence is a function of the duration of the stimulus at the time of the reinforcement.

Experimental evidence for the principle is meager and limited to classical conditioning situations. In a *delayed* conditioning study, Kappauf and Schlosberg (2) found that as the duration of the CS at the time of the US onset increased from a near-zero value, the amount of conditioning increased to a maximum at a critical duration value. With a further increase, the amount of conditioning decreased according to a simple decay function approaching an asymptotic value. It was upon the latter phase of the relationship that Hull based his principle of the dynamic state of the stimulus. Three other CR measures used by Kappauf and Schlosberg failed to show the relationship.

The initial growth phase of the relationship was explored precisely by Kimble (3), also in a classical conditioning study of the delayed type. The results indicated a negatively accelerated increase in the amount of conditioning with an increasing duration of the CS at the time of the US onset. Kimble did not explore the decay phase indicated by the Kappauf and Schlosberg experiment.

Although Hull applied the principle only to the decay phase, it is appropriately extended to both growth and decay phases. Accordingly, stimuli are increasingly less dynamic, i.e., have an increasingly lesser capacity for acquiring habit strength, as their durations at the time of the reinforcement deviate in either direction from some optional duration.

The present experiment investigates the validity of Hull's principle in a selective, trial-and-error discrimination problem rather than in a classical conditioning procedure. The discrimination problem required rats to make an appropriate T-maze response to a tone cue and the spatially opposite response to the absence of the tone. The duration of the tone was manipulated by having the onset precede the opportunity to make the choicepoint response by different intervals for each experimental group.

[1] This paper is part of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, in the Department of Psychology at the State University of Iowa. The writer expresses his appreciation to Professor Kenneth W. Spence for his critical advice throughout the course of the investigation.

Experimental Procedure

Subjects.—The *S*s were 56 female hooded rats from the colony of the Psychology Department of the State University of Iowa. The animals ranged in age from 90 to 120 days at the beginning of the experiment, and were distributed among the several experimental groups so as to equate the groups for age and littermate content.

Apparatus.—The apparatus consisted of two units, a straight runway used for preliminary training, and the single-unit T-maze. A floor plan of both units is presented in Fig. 1.

The end boxes consisted of two units, an outer shell, detachable from the arms of the maze, and a plywood insert which functioned as the goal box proper. The shell interiors were painted flat black, and the interior of the insert, a gloss white. The floor of the insert was differentiated from the floor of the shell by being covered with a rectangle of coarse hardware cloth. This insert-shell arrangement enabled *E* to change the side of the positive, white goal box (insert) according to a predetermined pattern by simply placing the insert into the appropriate shell. The spatially opposite end box was, then, the negative, or "incorrect," box. The negative, black shell was further differentiated from the positive, white insert in that the floor level of the latter was higher by the height of the insert floorboard, or approximately ¼ in. Thus, entry into the negative, black end box following an incorrect choice was distinguishable by visual, tactual, and proprioceptive differentials, from the entry into the positive, white goal box following a correct choice.

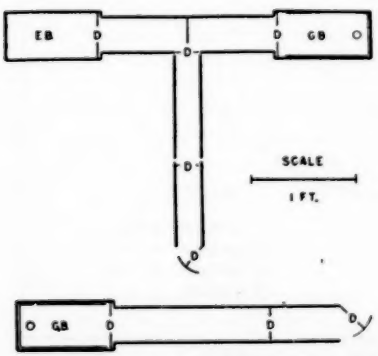

Fig. 1. The floorplan of the T-maze and the preliminary training runway. The letters *G.B.*, *E.B.*, and *D* refer to goal box, end box, and door, respectively.

The T-maze was constructed of pine wood pathways and covered with coarse hardware cloth. A manually operated guillotine-type door of masonite separated the starting-box chamber from the stem of the maze. A second guillotine-type door was located at the choicepoint end of the stem alley. This door was constructed in the form of an inverted T. It functioned as a barrier to the choicepoint, detailing *S*s in the stem delay chamber, and also served as a barrier to prevent retracing after a choice had been made into either arm of the maze.

The end boxes were separated from the maze proper by vertically sliding doors. In addition to the usual function of detaining the *S*s in the end boxes, the doors, when closed, prevented the interiors of the end boxes from functioning as differential cues to the correct choice.

A photoelectric beam was directed across the stem runway through two vertical slits in the runway walls. The beam was about ¼ in. from the starting-box door, so that entry into the stem runway immediately interrupted the beam.

The apparatus was so wired that on the "tone" trials, interruption of the photoelectric beam by *S* effected the onset of the tone. On "no-tone" trials, interruption of the beam had no effect. Termination of the tone was accomplished manually.

A Wright-Decoster loudspeaker was used to deliver the sound stimulus. It was suspended behind and above the choicepoint so that, tilted at a 35° angle, the center of the speaker was approximately 1 ft. from the choicepoint area. The auditory discriminandum was a 2,300-cycle tone generated by a Hewlett-Packard audio-oscillator, model 200 B. The intensity of the tone in the choicepoint area was 90 db above 10–16 w. per sq. cm.

A one-way vision screen separated the table maze from *E*. Two 200-w. lights, located 4 ft. above the center of each end box, provided the maze illumination. Except for the preliminary training, the experiment was conducted in a soundproof room.

Preliminary training.—All *S*s were placed on a 23-hr. food deprivation schedule five days before preliminary training. The schedule was maintained for the duration of the experiment. Preliminary training began in the straight runway. All *S*s were given five food-reinforced trials per day for six successive days. The purpose of this phase of the training was to acquaint *S*s with the goal box environment, and to establish a running-for-food habit of considerable strength.

Preliminary training was continued in the T-maze. All *S*s were given four forced trials per day to each goal box on the seventh and

eighth days. A single-alternation pattern was employed, and the pattern was reversed on the eighth day. The pattern was so initiated that the last forced trial of the preliminary training was to the side opposite from that which was correct on the first trial of the discrimination training. The purpose of this phase of preliminary training was to acquaint *S*s with the static conditions of the T-maze, and to provide an equal number of food-reinforced experiences with both goal boxes.

Discrimination training.—The discrimination problem required *S* to learn to choose one pathway when the tone condition prevailed, and to choose the spatially opposite pathway when the no-tone condition prevailed. Except for the *duration* of the tone or no-tone condition at the time of the opportunity to make the choicepoint response, the discrimination training procedure was identical for four of the five experimental groups. The tone durations for these four groups were, respectively, 1, 2, 5, and 15 sec. These durations were accomplished by physically delaying the *S*s of each group in the stem delay chamber for the appropriate interval before permitting a choicepoint response. The fifth experimental group encountered no physical delay. It operated as a control group against the possibility that the physical delay imposed upon the four main groups acted as a frustrative influence to retard learning. In this group, duration of the tone at the time of the choicepoint response opportunity was the time it took each *S* to run from the starting box to the choicepoint area. This temporal interval was determined approximately as .5 sec. for the group.

Running under a 23-hr. food deprivation condition, all *S*s were given ten trials per day, for which the white, food-containing goal box was located successively L (left), R (right), R, L, L, R, R, L, L, R. On alternate days, the mirror image of this pattern was employed. For half of each experimental group, the correct response to the tone condition was choice of the *right* pathway, while for the other half, the correct response to the tone was choice of the *left* pathway. A choice was defined as body-length entry into one arm of the maze. Correction was not permitted. A correct response to either the tone or no-tone condition was reinforced with a food pellet. The animals were given 15 sec. to consume the food pellet before being returned to the intertrial waiting cages.

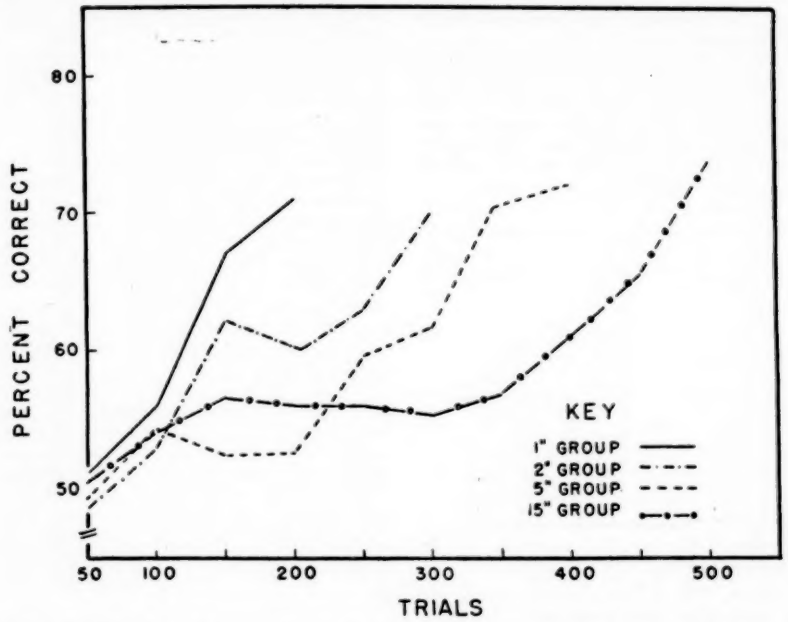

FIG. 2. The learning curves for the 1-, 2-, 5-, and 15-sec. tone-duration groups. The curves are limited to the number of trials for which the groups were intact.

Animals were detained in the end box after an incorrect choice for the same length of time. The intertrial interval was approximately 5 min.

For the tone trials, the onset of the tone occurred when an animal interrupted the photoelectric beam at the starting-box end of the stem delay chamber. For the no-tone trials, such interruption had no effect. Termination of the tone depended on the kind of response made. If incorrect, the tone was terminated only after *S* had been returned to the intertrial cage. If correct, the tone was terminated, for the initial 50 trials, at the moment *S* reached the white food cup, and, thereafter, immediately following the correct response.

Results and Discussion

The learning curves for the 1-, 2-, 5-, and 15-sec. tone groups are presented in Fig. 2 in terms of the percentage of correct responses for successive blocks of 50 trials. During the course of the experiment, it was found expedient to discontinue running *S*s that reached the learning criterion. Consequently, the learning curves are limited to the number of trials for which the experimental groups were intact. The curtailed family of curves, however, confirms Hull's principle of the dynamic state of a stimulus by indicating a decrease in the rate of learning with an increase in the duration of the tone at the time of the choicepoint response.

The mean and median numbers of trials to reach the 90 per cent learning criterion are given for the four main experimental groups in Table I. The *SD* values are also given. The gradient expressing the functional relationship between the rate of learning the discrimination and the duration of the cue stimulus at the time of the choicepoint response was derived from the median values.

TABLE I

Trials Required to Reach the 90 Per Cent Criterion

Measure	Tone Duration in Seconds			
	1	2	5	15
Median	170	250	370	500
Mean	190	274	385	517
SD	86.3	105.9	70.3	76.0

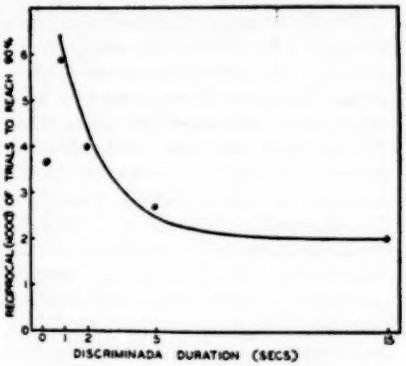

Fig. 3. The rate of learning a tone–no-tone discrimination as a function of the duration of the tone. The performance of the control group, which encountered no physical delay in the stem chamber, is shown by the point to the left of the gradient.

The gradient is presented in Fig. 3. The reciprocal of the median number of trials to reach the learning criterion multiplied by 1,000 is used as the index of rate of learning and is plotted against the stimulus duration values. It reveals a negatively accelerated decrease in the rate of learning with an increase in the duration of the cue stimulus. The precise relationship conforms to a simple decay function of the exponential type,[3] which agrees in form with the relationship derived by Hull (1, p. 179) from the Kappauf and Schlosberg data discussed earlier.

[3] The following equation was fitted to the data by a trial-and-error curve-fitting procedure:

$$Y = 4.40 \times 10^{-.20(T-.8)} + 2.0.$$

The value .8 in the equation represents an estimate of the optimal duration of the stimulus. The range within which the equation describes the relationship is from $T = .8$ sec. to $T = 15$ sec.

The functional relationship is theoretically interpreted as a decay in the capacity of a stimulus component to acquire an associative tendency to a response as the time of its action upon the involved receptor mechanism increases. Stated in terms of Hull's principle of the dynamic state of a stimulus, which it confirms, the relationship indicates a decrease in the magnitude of habit strength acquirable by a stimulus component with its increased duration at the time of an instrumental response and subsequent reinforcement. The principle is not at all unrelated to the common observation that steady-state stimuli of long duration lose their attention-provoking properties. In a learning context, this may be equivalent to loss by a stimulus of its capacity to enter into learned relationships.

The obtained gradient agrees with the Hull-derived Kappauf and Schlosberg gradient in two further respects: (1) There appears to be a critical cue stimulus duration for the maximal rate of learning; and (2) the gradient indicates a decay in the rate of learning to some asymptotic level which is approximately one-third of the initial height of the gradient.

With respect to the first observation, the empirical point to the far left of the gradient in Fig. 3 represents the rate of learning for the control group which encountered no physical delay in the stem chamber of the maze. The stimulus duration for the control group was determined as approximately .5 sec. The learning performance was inferior to both the 1- and 2-sec. stimulus duration groups. The significance of this is twofold.

First, it confirms a complex two-phase relationship between stimulus duration and the rate of learning. As the duration of the cue stimulus is increased from some near-zero value, the rate of learning increases to a maximal value at a critical duration, and as the duration is further increased, the rate of learning decreases as a simple decay function which approaches an asymptotic value. The solid-line gradient in Fig. 3 represents the latter phase.

Secondly, the inferior performance of the "no physical delay" control group assures that the variable underlying the gradient was not an increasingly severe disruptive effect resulting from longer durations of physical delay before the choicepoint response.[3] If this were the case, the control group should have been the superior group. This does not imply that physical restriction for periods longer than those used in the present experiment may not constitute a frustrating, even noxious, situation which would seriously interfere with the learning process. Such may well be the case. However, the action of such a secondary disruptive influence should be revealed by the nature of the empirical gradient at longer stimulus duration intervals. In the event of a sharp break from the present asymptotic level with longer stimulus durations, it would be reasonable to infer the introduction of such additional secondary factors.

With respect to the second observation drawn from the comparison of the Kappauf-Schlosberg gradient and the present one, no particular significance is attached to the fact

[3] An experiment by Muenzinger and Fletcher (4) provides evidence consistent with the interpretation that the short, physically enforced delays did not introduce disruptive emotional factors. They compared the rates of learning a black-white visual discrimination of a group physically delayed at the choicepoint with a group without physical delay. The short duration delay actually facilitated learning the discrimination problem.

that both gradients appear to decay to an asymptotic value approximately one-third the initial height of the gradient. That both decay to an asymptotic value, however, does have significant theoretical implications. If the gradients are interpreted as theoretical gradients reflecting a decline in the magnitude of a habit strength increment acquirable by stimulus components of increasing duration, then an asymptotic level of any value implies that no matter how long the duration of the cue stimulus, it will, other things equal, acquire an increment of habit strength. Furthermore, an increasing number of such long duration stimulus-response reinforcement sequences would imply a cumulative growth of the total effective habit strength available for the response. It is this theoretical interpretation which Hull has applied to the Kappauf and Schlosberg data, and which is supported by the present experiment. At the present time there is no adequate experimental evidence to the contrary. Even so, there is the possibility that the theoretical gradient decays to a zero magnitude.

There is one other aspect of the experiment which deserves comment. The marked similarity between the gradient obtained in this experiment and that obtained by Kappauf and Schlosberg in a classical conditioning design further emphasizes the fruitfulness of that aspect of Hull's theoretical system which attempts to derive the behavioral implications of complex learning problems from the primary principles of classical conditioning. Such agreement argues against the interpretation that classical conditioning involves a learning process fundamentally different from that involved in selective, trial-and-error learning.

Summary

The experiment was designed to examine the validity of Hull's principle of the dynamic state of a stimulus, which states, in effect, that the capacity of a stimulus component to acquire habit strength decreases as the time of the action of the stimulus upon the receptor mechanism increases. Specifically, the experiment investigated the nature of the functional relation between the duration of a tone at the time of the opportunity to make a choicepoint response, and that rate of learning a tone–no-tone discrimination. Four groups of *S*s were trained to respond differentially to the presence or absence of a tone at the choicepoint of a T-maze, after enforced exposure to the tone for 1, 2, 5, and 15 sec.

A plot of an index of the rate of learning against the duration of the discriminandum revealed a gradient which shows a simple decay from a maximum value to an asymptotic limit above zero. The gradient is interpreted as confirming Hull's principle of the dynamic state of the stimulus.

(Manuscript received June 24, 1950)

References

1. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
2. Kappauf, W. E., and Schlosberg, H. Conditioned responses in the white rat: III. Conditioning as a function of the length of the period of delay. *J. genet. Psychol.*, 1937, **50**, 27–45.
3. Kimble, G. Conditioning as a function of the time between conditioned and unconditioned stimuli. *J. exp. Psychol.*, 1947, **37**, 1–15.
4. Muenzinger, K. F., and Fletcher, F. M. Motivation in learning: VII. The effect of an enforced delay at the point of choice in the visual discrimination habit. *J. comp. Psychol.*, 1930, **23**, 383–392.

THE RÔLE OF IRRELEVANT DRIVE STIMULI IN LEARNING THEORY[1]

BY M. U. ENINGER

Carnegie Institute of Technology

A current controversy within the framework of Hull's (3) theoretical system concerns the question of whether irrelevant drive stimuli[2] have the capacity to acquire associative strength as do relevant drive stimuli and external stimuli. Hull implies that they do, that relevancy does not influence the capacity of drive stimuli to enter into associative connections. A recent hypothesis by Kendler (4), termed *selective association*, maintains the contrary. Only relevant drive stimuli, according to Kendler, become connected to a rewarded response. The hypothesis is based on the finding that rats, after having been run while both hungry *and* thirsty in a T-maze in which one arm led to food and the other to water, were subsequently able to choose the appropriate side when either hungry or thirsty. These results are not predictable on the basis of the assumption that relevant and irrelevant drive stimuli have an *equal* capacity for associative strength. They are predictable, however, from the assumption Kendler makes, namely, that only relevant drive stimuli acquire associative strength. A theoretical picture of the conflicting assumptions in the context of Kendler's experiment, is presented in Fig. 1.

More recently, Amsel (1) provided evidence contradictory to the selective association hypothesis. This investigator required rats to learn to choose one arm of a T-maze to reach an anxiety-reducing goal box when under one irrelevant need condition, and the alternative arm under a different irrelevant need condition. Because the irrelevant need conditions alone provided the basis for the discrimination, evidence of learning would disprove Kendler's hypothesis. Only two of Amsel's ten *S*s, however, reached the learning criterion before training was discontinued. Amsel explains the slowness of learning by maintaining that both relevant and irrelevant drive stimuli acquire associative strength only with great difficulty. This is so because the

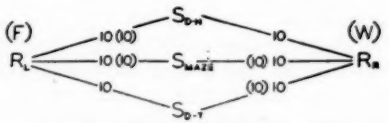

FIG. 1. The diagram represents the theoretical balance of associative strength presumed to be operating in Kendler's *S*s after training. The open values represent assumed magnitudes of associative strength on the assumption that relevancy does not affect associative capacity. The bracketed values represent assumed magnitudes according to the selective association hypothesis. S_M, S_{D-H}, and S_{D-T} represent maze stimuli, and hunger and thirst drive stimuli, respectively. *F* and *W* represent food and water reinforcements which followed R_{left} and R_{right}, respectively. The theoretical balance of associative strength is shifted in favor of the appropriate response when a single need condition prevails, only if the "selective association" assumption is made.

[1] This paper is based on a thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Psychology at the State University of Iowa. The writer expresses his appreciation to Professor Kenneth W. Spence for his critical advice throughout the course of the investigation. The author is solely responsible for the interpretations made.

[2] Drive stimuli are said to be "relevant" when they are reduced in intensity by the goal object in the experimental learning problem. Drive stimuli which are not so reduced are termed "irrelevant."

onset of such stimuli precedes the experimental situation by many hours as a rule. And according to the principle of the dynamic state of stimuli proposed by Hull (3, p. 207), and recently confirmed in discrimination learning (2), the capacity of a stimulus to acquire associative value decreases with increasing duration of the stimulus at the time of reinforcement.

The present experiment attempts to provide more definitive evidence that irrelevant drive stimuli do enter into associative connections. The specific problem is to determine whether *S*s with the *same* irrelevant drive condition present during the extinction of a response as was present during the learning of the response show greater response stability than a comparable group which undergoes extinction under a *different* irrelevant drive condition from that present during the learning of the response. If the irrelevant drive stimuli do acquire associative connections to the response, the second group would have less total associative strength involving the response during extinction, and should demonstrate less response stability.

Experimental Procedure

Subjects.—The *S*s were 40 naive albino and hooded rats, of both sexes, from the colony maintained by the Department of Psychology at the State University of Iowa. They ranged in age from 75 to 120 days at the beginning of the experiment. The *S*s were assigned to the experimental groups so as to equate the groups for age, strain, and sex content.

Apparatus.—A single-unit T-maze was used. The starting box and runway of the maze were 2½ in. wide and 3½ in. high, and were covered with removable Plexiglass panels. The floor of the maze was a grid of parallel brass rods, $\frac{1}{16}$ in. in diameter and spaced $\frac{1}{16}$ in. apart. It was connected with a Muenzinger and Walz (6) type shock apparatus and could be activated by a foot pedal.

TABLE I

Irrelevant Drive Conditions Imposed on *S*s during Training and Extinction

Training	Extinction	
	Hunger	Thirst
Hunger	(H-H)	(H-T)
Thirst	(T-H)	(T-T)

Attached to both arms of the T-maze were removable goal boxes. They were constructed of unpainted plywood and covered with a panel of hardware cloth. The goal boxes were identical in physical dimensions and appearance.

The doors throughout the maze were vertically sliding panels of masonite. They were located at strategic points to prevent retracing. The maze was illuminated by two 40-w. lights, located approximately 2½ ft. above the goal boxes. A one-way vision screen separated the tabled maze from *E*.

Procedure.—Four experimental groups were used, and the irrelevant drive conditions which define the groups are summarized in Table I. Two of the groups (H-H and T-T) underwent extinction under the same irrelevant drive condition present during learning, and two groups (H-T and T-H) underwent extinction under a different irrelevant drive condition from that present during learning. Deprivation for 22 hr. was used for both irrelevant hunger and thirst needs. The number of *S*s in each group was ten.

Ten days before training was initiated, the *S*s were put on a fixed feeding schedule which preceded by 2 hr. the time of the experimental sessions. The *S*s were handled frequently by *E* during this period. On the third day before training, *S*s were subjected to the irrelevant need deprivation conditions to which they were assigned for the training series.

Training.—All *S*s received 10 escape-from-shock reinforced trials per day for five days, a total of 50 training trials. The *S* was placed in the starting box with the grid floor activated. Choice of either the left or right goal box on the initial run designated the goal box into which escape was permitted on subsequent training trials. Thus, the goal box of the initial choice was the "correct" goal box. An incorrect choice on subsequent trials required the *S* to correct its run in order to escape shock. Each *S* was permitted 30 sec. in the goal box, and the intertrial interval for any one *S* varied between 10 and 15 min. Neither food nor water was

presented in the goal box, the reinforcement being solely escape from shock.

Extinction.—Beginning on the day following the last training trial, all *Ss* were given five no-shock or extinction trials per day for four days. It was assumed that the anxiety aroused by the maze environment would motivate the *Ss* to seek escape, and such was the case.

For the initial ten extinction trials, escape was permitted only into the goal box designated "correct" on the basis of an *S*'s initial run. Choice of the opposite arm of the maze required correction in order to escape the anxiety-inducing environment. For the last ten extinction trials, escape was permitted into either goal box. Throughout the extinction trials, however, the "correct" response was the escape response learned during training. Thus, the extinction procedure is a measure of the stability of the learned escape response under the four combinations of irrelevant need conditions.

Results

The mean and median numbers of correct responses made by the four experimental groups during the extinction series are shown in Table II, together with the mean and median numbers of correct responses for the combined groups, H-H and T-T, and the combined groups, H-T and T-H. The mean number of correct responses for *Ss* extinguished under the *same* irrelevant drive stimulus condition as imposed during training is 12.64; for those extinguished under *different* irrelevant drive conditions, the mean value is 8.88. A *U*-test (**5**), which makes no assumption about the nature of the distributions from which the means are calculated, indicates the difference between the means for the combined groups to be significant at between the .02 and .01 levels of confidence ($Z = 2.26$).

TABLE II

Correct Responses During Extinction Trials When Irrelevant Drives Were the Same or Different During Learning and Extinction

	Same Irrelevant Drives			Different Irrelevant Drives		
	T-T	H-H	Combined	T-H	H-T	Combined
Mean	12.83	12.46	12.64	8.50	9.23	8.88
Median	14.00	13.00	14.00	8.50	9.00	9.00

The results are clearly contradictory to Kendler's principle that irrelevant drive stimuli do not have the capacity for associative connections. If, during training, the irrelevant drive stimuli acquired no associative connection to the escape response, no difference in resistance to extinction by the two combined groups could be predicted. The contrary assumption, however, predicts the difference in response stability as it occurred.

If, on the basis of Amsel's and the present results, Kendler's hypothesis of selective association is rejected, how may his findings be explained? An alternative theoretical account, which involves the concept of the fractional anticipatory goal response mechanism, is provided by Kendler (**4**) and supported by Amsel (**1**). A simpler theoretical account is given by the assumption that irrelevant drive stimuli do acquire increments of associative strength *relating them to overt responses*, but of a lesser magnitude than those acquired by relevant drive stimuli.

Summary

The present experiment tested the hypothesis that irrelevant drive stimuli have no associative capacity. White rats that had the same irrelevant need condition present during the extinction of a maze response as was present during the learning of the response showed greater response stability than a comparable group which underwent extinction under a different irrelevant need condition from that present during the response

acquisition. The results were interpreted as contradictory to Kendler's hypothesis of selective association.

An alternative hypothesis was offered to account for Kendler's findings, namely that relevant drive stimuli acquire a greater increment of associative strength per reinforcement than irrelevant drive stimuli.

(Manuscript received June 24, 1950)

References

1. Amsel, A. Selective association and the anticipatory goal response mechanism as explanatory concepts in learning theory. *J. exp. Psychol.*, 1949, **39**, 785–799.
2. Eninger, M. U. The rate of learning a tone–no-tone discrimination as a function of the tone duration at the time of the choicepoint response. *J. exp. Psychol.*, 1951, **41**, 440–445.
3. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
4. Kendler, H. H. The influence of simultaneous hunger and thirst drives upon the learning of two opposed spatial responses of the white rat. *J. exp. Psychol.*, 1946, **36**, 212–220.
5. Mann, H. B., and Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Statist.*, 1947, **18**, 50–60.
6. Muenzinger, K. F., and Walz, F. C. An examination of electrical-current stabilizing devices for psychological experiments. *J. gen. Psychol.*, 1934, **10**, 477–482.

THE EFFECT OF TRAINING PROCEDURES ON THE RELATIVE STRENGTH OF PLACE AND DIRECTION DISPOSITIONS

BY FREDERICK R. FOSMIRE

Montana State University

AND

W. LYNN BROWN

The University of Texas[1]

The series of studies projected by Tolman, Ritchie, and Kalish on "spatial learning" postulate that development of a "cognitive map" by the rat accompanies maze experience. This "cognitive map" enables the rat, in future encounters with a particular maze, to go to the place in the room where food has previously been experienced. They call the disposition to go to the place a "place disposition." In reply to the Blodgett and McCutchan criticisms (**1**, **2**), Ritchie, Aeschliman, and Peirce (**7**) try to state with extreme clarity what was meant in an earlier article (**9**) by the expressions "place learning," "response learning," "place disposition," and "response disposition." They point out that in the earlier paper Tolman, Ritchie, and Kalish (**9**) "confuse the notions of place-disposition with the notion of place-learning in such a way that it made it appear that place-learning was defined as the acquisition of a place-disposition" (**7**, p. 74). The same general statement would apply to the term "response learning." Ritchie, Aeschliman, and Peirce (**7**, p. 73) also say that "an animal has a *place-disposition*" when "this animal, when placed in a *new* situation (one in which his spatial relation to the discriminable extra-maze cues is new to him), will choose a path which takes him towards the goal place, or in the direction of the goal place." In an earlier article Ritchie (**6**) distinguishes between *place orientation* and *direction orientation* as two kinds of place dispositions.

Blodgett, McCutchan, and Mathews (**3**), referring to the first Tolman, Ritchie, and Kalish article (**8**), present an experimental arrangement in which response disposition and place disposition would be opposed. One of the conclusions drawn from their experimentation is that *S*s, in learning simple **T**-mazes, tend to acquire direction dispositions. They state: "We find no evidence that our subjects have acquired a place-disposition" (**3**, p. 805).

Gentry, Brown, and Kaplan (**4**) and Gentry, Brown, and Lee (**5**) present results which fail to support the spatial-learning hypothesis projected by Tolman, Ritchie, and Kalish (**8**). Gentry, Brown, and Kaplan (**4**, pp. 321–322) conclude: "Since our results for both the animals with training and without training are in part explainable in terms of an artifact of the experimental apparatus (the presence of alley CD), we wonder if their results are not explainable in terms of some artifact of their experimental set-up, and hence if there is any need for them to appeal to the senior

[1] The cost of this investigation was met in part by a grant to the senior author by the Research Institute of The University of Texas.

author's expectancy theory in its inferential form."

It occurred to the writers that acquisition of a disposition (place or direction) might be in part a function of the training procedures utilized. In this experiment we used a procedure not unlike that of Blodgett, McCutchan, and Mathews (3), the major difference between the two procedures being that our *S*s would not be required to demonstrate a disposition except on the test runs. We thought that such a change in procedure might favor the acquisition of a place disposition instead of a direction disposition as Blodgett, McCutchan, and Mathews found (3).

Problem

The present experiment was designed to test the effect of training procedures on the relative strength of place and direction dispositions by forcing *S* to choose between two opposing responses: a place response and a direction response after training on a straight runway.

Experimental Procedure

Subjects.—Twenty-nine experimentally naive *S*s were used in this experiment. They were divided into groups designated I, II, and III, containing 10, 9, and 10 animals, respectively. All *S*s were descendants of the Wistar strain reared in the Comparative Psychology Laboratory of the University of Texas. They were approximately 120 days old at the beginning of the experiment.

Apparatus.—The maze setup is presented in Fig. 1. The runways were 1½ in. wide and 15 in. above the floor and were painted flat black. The main runway was 7 ft. long and had a removable brass food cup countersunk at one end. Only one main runway was in the experimental room during any training period or during test runs. All preliminary training was done on the main runways. A 24-in. runway was used on test runs in order to make a T-maze. Only one T-maze was set up in the room at a time during testing.

Feeding.—The *S*s were placed on a 24-hr. feeding schedule for seven days before training began. The diet used during this experiment was five parts (by weight) of yellow corn meal, one part of linseed meal, one part of powdered buttermilk, one part of ground alfalfa. About 50 gm. of salt and 50 gm. of calcium carbonate were added for each five pounds of corn meal. The mixture was moistened with water. The *S*s were permitted to eat for about 30 sec. after each training run and for 15 sec. after each test run.

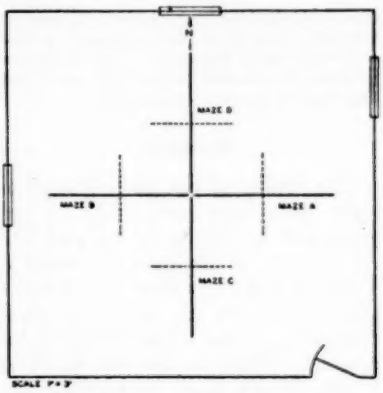

Fig. 1. Room and mazes

Training.—The *S*s were run four times a day between 1 and 4 P.M. for seven days. The *S*s in Group I were run on the main runway of Maze A from east to west. Group II *S*s were given two runs a day on the main runway of Maze A (east to west) and two runs daily on the main runway of Maze B (west to east). Group III *S*s were given two runs daily on the main runway of Maze A (east to west) and two runs daily on the main runway of Maze C (south to north). The order of runs for Groups II and III was varied in a haphazard manner. All *S*s were rewarded at the *same place* in the room.

Test procedure.—On the eighth day, after 28 training runs for each *S*, a 24-in. starting pathway was introduced, making a T-maze for each test position. Two test runs were given on each maze (A, B, C, and D) for each *S* in the three groups, one from each side of the main pathway. On some of these test runs *S* had to choose between running in the direction in which it had been trained to run and running to the place in the room to which it had been trained to run. That is, when *S*s in Group I were tested on Maze B, two possible cues as to where the food was located were opposed. If *S* had been responding during training on the basis of a direction disposition, it would be expected to

choose to run from east to west, an incorrect response. But if *S* had acquired a place disposition during training, it would be expected to inhibit the tendency to run from east to west and, instead, run from west to east.

Results

The results of the three groups on eight test runs (see Fig. 1) are presented in Tables I, II, and III. An analysis of the data presented in Table II will show that the *Ss* had a slight tendency to adopt a place disposition rather than direction disposition. Only on the test on Maze B did the *Ss* favor a direction disposition response. The percentage of direction disposition responses on Maze B (63.7 per cent) could be expected on the basis of chance alone between 10 and 20 per cent of the time.

It was also found that differences in training procedures influence disposition. In general, the training procedure used with Group II was more conducive to the establishment of place disposition than was the procedure used with Group I. Likewise, the procedure used with Group III was more conducive to the establishment of place disposition than was the procedure used with Group II.

Although in this experiment there was a slight tendency for *Ss* to utilize place disposition more than direction disposition, there were considerable individual differences. One *S* in Group II (see Table III) adopted a place disposition to a degree that correct response was made on all eight test runs. In the three groups there were 18 *Ss* that favored place disposition to some extent. The remainder of the *Ss* either did not demonstrate a preference or had a direction disposition.

The results also suggest that when certain relationships between maze and physical surroundings exist, one disposition will dominate, but when other relationships exist, another disposition may dominate. Consider the maze setup in the experimental room. When *Ss* are trained on Maze A, they are running from a relatively homogeneous wall, with a window to one side, directly toward a window. When they are trained on Maze B, they are running from a window

TABLE I

Per Cent Correct Runs on Test Trials

	Group I	Group II	Group III
When running west	70	67	60
When running east	30	39	40
When running north	65	67	75
When running south	55	56	60

TABLE II

Significance of Place Performance Tendency on Test Trials

	Per Cent Place Performance (Groups I, II, and III)	Chi-Square	*P*
When running west	69.5	4.4	.05
When running east	36.2	2.2	.10–.20
When running north	69	4.2	.05
When running south	56.9	0.6	—

TABLE III

Individual Differences in Place Performance Tendencies During Test Trials

Per Cent Place Performance	Number of Animals		
	Group I	Group II	Group III
100.0		1	
75.0		1	1
62.5	6	3	6
50.0	2	1	2
37.5	2	2	1
25.0		1	
Total	10	9	10

toward a relatively homogeneous wall with a window to one side. When they are trained on Maze C, they are running from a homogeneous wall directly toward a window. The *Ss* in Group I were trained only on Maze A but were tested on Mazes A, B, C, and D. The *Ss* in Group II were trained on Mazes A and B and were tested on all four mazes. The *Ss* in Group III were trained on Mazes A and C and tested on all four mazes. The *Ss* in Group I (trained only on Maze A) showed greater direction disposition than either group. The *Ss* in Group III (trained on Mazes A and C) showed greater place disposition than either group. The *Ss* in Group I had a very strong tendency to run toward the bright window (west) when running Mazes A and B on the test runs.

Discussion

The present experiment was designed to study the development of place disposition and direction disposition as a result of training procedures. The results indicate a slight tendency for *Ss* to adopt a place disposition rather than a direction disposition. There were also individual variations among *Ss* as to which disposition was adopted, there being more that favored place disposition than direction disposition.

In addition to the individual differences in disposition, there are differences attributable to training procedures utilized. The training procedure utilized with Group III was more conducive to establishment of place disposition than was that utilized with Group II, whereas Group II training procedure developed place disposition slightly more readily than Group I procedure. This can probably be explained in terms of ambiguity of the training situation. Since the *Ss* in Group I were trained on the main runway of Maze A only, they probably did not learn to make a distinction between place and direction (if Group I *Ss* could "know" a place response from a direction response). This ambiguity is not present in the Group II training procedure and, theoretically, if there were nothing about the physical surroundings to influence responses systematically on the test runs, we should expect this training procedure to yield a place disposition more often than the Group I training procedure. That is, if there were no outstanding differences between the east and west walls of the experimental room, we would expect the Group II training procedure to be much superior to the Group I procedure as far as establishment of place disposition is concerned. The training procedure for Group III may or may not be more conducive than Group II procedure to the establishment of place disposition in rate, depending on whether or not it is functionally a single problem or two separate concurrent problems. If, as far as the *S* is concerned, the procedure utilized for Group III is considered as training for a single problem, it will probably yield similar incidences of place disposition. If it is viewed as two training procedures for two different problems, it will probably yield a lower incidence of place disposition.

Summary and Conclusions

1. Three groups of 10, 9, and 10 *Ss* were subjected to three different training procedures designed to make possible to the *S* adoption of a place disposition or direction disposition.

2. After 28 training runs *Ss* were each given eight test runs. It was

found that when place-disposition responses and direction-disposition responses are opposed, more *Ss* respond on the basis of place disposition than on the basis of direction disposition.

3. Individual differences in acquisition of the two kinds of dispositions were found.

4. Attention is called to the possibility that acquisition of a particular disposition is a function of the training procedure used and of any striking characteristics of the physical surroundings.

(Manuscript received for immediate publication February 17, 1951)

References

1. Blodgett, H. C., and McCutchan, K. Place versus response learning in the simple T-maze. *J. exp. Psychol.*, 1947, **37**, 412–422.
2. Blodgett, H. C., and McCutchan, K. The relative strength of place and response learning in the T-maze. *J. comp. physiol. Psychol.*, 1948, **41**, 17–24.
3. Blodgett, H. C., McCutchan, K., and Mathews, R. Spatial learning in the T-maze: The influence of direction, turn, and food location. *J. exp. Psychol.*, 1949, **39**, 800–809.
4. Gentry, G., Brown, W. L., and Kaplan, S. J. An experimental analysis of the spatial location hypothesis in learning. *J. comp. physiol. Psychol.*, 1947, **40**, 309–322.
5. Gentry, G., Brown, W. L., and Lee, H. Spatial location in the learning of a multiple-T maze. *J. comp. physiol. Psychol.*, 1948, **41**, 312–318.
6. Ritchie, B. F. Studies in spatial learning. VI. Place orientation vs. direction orientation. *J. exp. Psychol.*, 1948, **38**, 659–669.
7. Ritchie, B. F., Aeschliman, B., and Peirce, P. Studies in spatial learning. VIII. Place performance and the acquisition of place dispositions. *J. comp. physiol. Psychol.*, 1950, **43**, 73–85.
8. Tolman, E. C., Ritchie, B. F., and Kalish, D. Studies in spatial learning. I. Orientation and the short-cut. *J. exp. Psychol.*, 1946, **36**, 13–24.
9. Tolman, E. C., Ritchie, B. F., and Kalish, D. Studies in spatial learning. II. Place learning vs. response learning. *J. exp. Psychol.*, 1946, **36**, 221–229.

USE OF SPECTRAL HUE-INVARIANT LOCI FOR THE SPECIFICATION OF WHITE STIMULI

BY DOROTHEA JAMESON AND LEO M. HURVICH

Color Control Division, Eastman Kodak Company

The concept *white* or *neutral* occupies a fundamental place in visual theory and in theoretically oriented visual research. For example: complementary colors are determined by equating a mixture of two spectral components to a "neutral" standard; saturation scales are based on the amount of "white" light required to extinguish a specific hue sensation; changes in visual sensitivity are evaluated from data obtained following adaptation to monochromatic and "white" stimuli; theoretical inferences are drawn from measured variations in hue when "white" light is added to a spectrally homogeneous stimulus. Since the conclusions, in each of these instances, depend, for their validity, on the nature of the "white" employed in the experiment, it is unfortunate that the "white" standards used differ considerably from one experiment to the next.

The term "white light" has come to be used as a generic term for a group of physical stimuli whose spectral distributions are as different as those of incandescent indoor illumination and the light from an overcast sky. This usage derives mainly from the operation of approximate color constancy, and for this reason it is valid in the everyday situation. Because of the extensive adaptive capacity of the visual mechanism, lights which are widely different physically tend to be perceptually equivalent, and modifications in the physiological system which make possible this perceptual equivalence are, in this context, irrelevant. In visual research, however, such physiological modifications cannot properly be ignored. When the term "white" is used as a category for visual stimuli, it not only describes the sensation evoked, it also implies certain specific assumptions concerning the physiological processes which mediate the white sensation. Moreover, theoretical deductions from the results of experiments employing "white" stimuli depend on the correctness of these implied postulates.

When a visual stimulus is categorized as white, it is assumed that such a stimulus excites the different physiological chromatic processes to an equal degree and that consequently its continued action produces no change in the relative excitability or responsiveness of the different mechanisms with respect to each other. On the basis of the perceptual criterion alone, it is practically impossible to test the validity of these basic assumptions, even under controlled laboratory conditions. A wide variety of physical stimuli are found to appear white, even when the visual system is in an initially neutral state, and the stimulus exposures are of only a few seconds' duration (**8**, **9**). Furthermore, controlled variations in the stimulus duration make it clear that even for short exposures the perceptual equivalence of the different stimuli may result, not from physiological responses to these stimuli which are initially identical and continue to be so, but rather from very rapid sensitivity changes in the differently excited color receptors during the action of the primary stimulus. For example, a stimulus may appear

slightly blue during a 1-sec. exposure, yet it rapidly desaturates and appears neutral at the end of a 5-sec. viewing period. The initially stronger excitation of the blue chromatic process obviously leads to a rapid desensitization of this process, and the continuing action of the stronger blue stimulus value on the less responsive blue mechanism now brings about excitations which are equal to those in the other chromatic mechanisms. Similarly, a stimulus may appear slightly yellow during a 1-sec. exposure, yet rapidly desaturate and become neutral when viewed for 5 sec., indicating in this case a rapid loss in the relative sensitivity of the yellow process, etc. Perceptual equality is achieved in all cases, despite differences in the stimulus chromaticities and differences in the initial chromatic excitations, only as a consequence of these rapid and compensatory adaptive changes. To eliminate such adaptive variations completely would require the use of infinitely short exposures, and hence the perceptual method for the selection of a white stimulus becomes impractical.

Since the perceptual criterion alone is inadequate, have we any other means for determining those stimuli which excite the chromatic processes equally in the non-fatigued eye and also maintain the system in a physiologically neutral or balanced state? The problem now becomes one of measuring chromatic sensitivity changes which take place during adaptation.

Ordinarily, visual sensitivity changes in chromatic adaptation are studied and evaluated in terms of changes in perceived colors. Sometimes, these effects are determined simply by reporting changes in the appearance of a given stimulus with its continued action (**12**). Quantitative measurements are usually obtained by a comparative technique in which equations are made between a test stimulus acting on a retinal area adapted to the chromaticity in question and a variable stimulus acting on a separate retinal area (either in the same or the other eye) whose adaptive state is assumed to be known. In some experiments the standard reference state is that of dark adaptation (**16**); in others it is that of bright adaptation to a light presumed to be white or neutral (**1**). The particular conditions chosen depend, of course, upon the answers sought in the given experiment.

To answer the questions we have posed concerning the balance of chromatic sensitivities, neither of these quantitative techniques is adequate. If the bright-adapted condition provides the standard reference state, the measured chromatic sensitivity changes will depend on the quality of the standard bright-adapting stimulus. Without some independent and prior guarantee that this stimulus is, in fact, truly neutral, an inescapable circularity inheres in the situation. The dark-adapted condition as the standard reference state presents other difficulties, however. A color equation made initially with both retinal areas in the standard, dark-adapted condition is compared with subsequent equations following pre-exposure of the test area to the adapting stimulus. Apart from any changes in relative chromatic sensitivities which may be induced by the adapting stimulus, the test area necessarily undergoes changes due to simple bright adaptation, i.e., a reduction in white or general light sensitivity. Equations made to the test stimulus must now reflect the decreased brightness and increased saturation which result from the

bright adaptation per se (6, p. 170; 7), and these effects are not readily isolated from changes brought about by alterations in the balance of chromatic sensitivities. Clearly, some functional property of the visual mechanism must be found which is dependent on the physiological balance of chromatic sensitivities and yet independent of changes in general light sensitivity or simple bright adaptation.

As such a functional index, Tschermak has employed the locations of the psychologically pure or unique hues in the spectrum (13; 15, p. 65). The usefulness of this index depends on the validity of the basic assumption that the wave lengths which evoke a single-hued sensation—i.e., pure blue, pure green, and pure yellow—will necessarily undergo a shift with changes in the relative chromatic sensitivities of the visual system, but will be unchanged with variations in simple bright adaptation. There seems to be some doubt, however, whether the wave lengths associated with the pure hues in the spectrum are constant or whether they differ for different brightness levels (10, p. 160). If the wave lengths of the pure hues are dependent on brightness, then bright adaptation per se might also be expected to introduce shifts in the spectral locations of these hues even in the absence of any change in the chromatic balance of the system. It is agreed, however, that there are three spectral stimuli which are invariant in hue with changes in intensity or brightness,[1] whether or not these hues are identical with the psychological primaries or unique hues. Consequently, the locations of the invariant hues in the spectrum were employed in the present investigation.[2]

The particular wave lengths of the three spectral invariant hues depend, of course, upon the spectral response properties of the visual system, and any changes in these properties must consequently be reflected in changes in the invariant loci. The existence of the spectral hues which are invariant with intensity has long been recognized, and the wave lengths of these stimuli have been measured by several investigators (3, 4, 11).[3] The locations of these hues have not, however, been used as indices of the functional state of the visual mechanism, and such an application would indeed be impractical using the methods which have been employed in the past for determining the invariant loci in the spectrum. The most extensive quantitative study of spectral hue and intensity is that of Purdy (11), who measured the change in hue of spectral stimuli for different pairs of intensities and, thus, also determined the wave lengths of the three spectral stimuli which do not change in hue with variation in intensity. The procedure by which his determinations were made was necessarily a long and tedious one, since it involved making hue equations for narrowly separated bands across the spectrum for each of six paired intensities. For

[1] There is, of course, a fourth invariant hue which is extraspectral, formed by mixing stimuli from the long and short wave length regions of the spectrum.

[2] Other techniques have been suggested which are concerned specifically with evaluating the effective neutrality of a stimulus acting on the visual system. One such method is Hering's "dimming" technique (2). Forbes (5) has suggested that the white standard be determined on the basis of the perceptual quality of the "Hering after-image." Neither of these proposals, however, provides a means for quantitative evaluation of the effects of different stimuli which are perceptually neutral.

[3] Three spectral hues have also been shown to be invariant both with retinal position and adaptation (14, p. 345–346).

the purposes of our application, however, the loci of the invariant hues alone are of interest, and the amount of the hue shift in other spectral regions does not concern us. Hence, the procedure employed to determine the wave lengths of the invariant stimuli can be considerably simplified as compared with the Purdy technique.

Experimental Technique

Apparatus.—Homogeneous spectral light of continuously variable wave length was provided by a Farrand monochromator (Model 300 Vis.). The light source was a 6-v. Nitra lamp, and its intensity was controlled at the entrance slit of the monochromator by means of a variable Wratten neutral wedge. At the fixed slit width of .1 mm., the wave length band in the blue region was about 1.3 millimicrons and in the yellow region, 2.3 millimicrons. A short-focus monocular telescope was used to view the exit slit of the monochromator. The upper half of the slit was covered with a Wratten neutral gelatin filter of an average density of 1.0, and *O* accordingly saw a magnified slit image of a single spectral region at two different brightness levels. As seen through the telescope, the apparent width of the slit was 1°12′, and the apparent length of each half slit was 16°. Fixation was maintained on the mid-point.

The test stimuli to which the eye was adapted in the various experiments were provided by an independent optical system. The light source was a 6-v. flat ribbon filament lamp, run at a constant voltage to maintain the source output at a calibrated color temperature of 2842° Kelvin. The diverging light from the source flooded the field of view of the ocular through which *O* viewed the adapting stimuli. The intensity of the beam was controlled by means of an iris diaphragm placed directly in front of the source. Wratten color correction filters inserted in the beam were used to vary the spectral distribution of the source, and thus the effective color temperature of the adapting stimulus. The angular subtense of the adapting field was about 47° at *O*'s eye. The intensity of the beam was adjusted to provide a constant luminance of about 25 millilamberts for all color temperatures of the adapting stimulus. The approximate color temperatures of the stimuli used were: 10,000° K, 7500° K, 5500° K, 4800° K, and 4300° K. Since a completely homogeneous adapting field was desired for these experiments, no fixation point was used. *O* was instructed simply to fixate approximately in the center of the field of view. The angular subtense of the adapting field was large in comparison to that of the spectral test stimuli; thus, small fluctuations in *O*'s fixation during adaptation probably had little or no influence upon the experimental results.

Procedure.—In brief, the experimental procedure consisted of determining, first, the wave lengths of the three invariant hues in the spectrum when *O* was in a neutral state of adaptation, and then redetermining these loci after controlled adaptations to various stimuli. To determine the wave length of an invariant hue, *O* scanned a given region of the spectrum by rotating the wave length drum of the monochromator, thus sweeping the spectrum past the fixed exit slit and producing a continuous variation in the wave length of the stimulus which filled both the brighter and less bright halves of the rectangular field of view. A wave length which is invariant in hue is readily detected by such a technique, since variations in hue with change in wave length take place at different rates for the two brightness levels. Thus, a change in wave length which produces only a small hue change in the brighter field half may produce a much more pronounced shift in the less bright half, and vice versa. Moreover, the resultant hue differences between the brighter and less bright field halves occur in the opposite sense on the two sides of a hue invariant locus. In the blue region of the spectrum, for example, a wave length somewhat higher than that of the blue invariant locus will appear predominantly blue in the brighter field half, but tinged with green in the less bright field half. A wave length slightly lower than the blue invariant locus, on the other hand, will again appear predominantly blue in the brighter field half, but now the less bright field half will appear tinged with violet rather than with green. In the region of the green invariant locus, the brighter field half appears more yellow than the less bright one for longer wave lengths than that of the green invariant hue, and for shorter wave lengths it appears more blue than the less bright field half. In the yellow region the brighter field half again appears predominantly yellow at both higher and lower wave lengths than the invariant locus, while the less bright field half is tinged with red or green, respectively. Thus, it is a fairly easy discriminatory judgment to single out each of the three transition regions where there is no difference in hue with a difference in intensity.[4]

[4] The technique used for determining the invariant loci is essentially a refinement of that employed by Exner (4), who simply compared two complete spectra of different intensities, and

It is essential that some specifiable condition of the visual system be established as a standard reference state. We designate this reference condition as the "neutral state." A series of preliminary observations indicated that a 10-min. period of preliminary dark adaptation is adequate for the visual system to recover from the effects of previous uncontrolled chromatic stimulation and to reach this neutral equilibrium state. The determination was made empirically in the following manner. *O* simply entered the darkroom in which the experiment was conducted, and without any controlled preadaptation period, immediately made a determination of the wave length of one of the invariant hues. Following a 2-min. period in the darkroom with *O*'s eyes completely shielded from any external light, another determination of the same locus was made. Repeated determinations were made at 2-min. intervals for approximately 30 min. It was found that the measured wave length of a given invariant hue changed only during the first few minutes in the darkroom. After 10 min. or less in the dark no further systematic variation occurred in the measured invariant locus, and it was assumed that this period of light exclusion was adequate to restore the visual system to a neutral equilibrium condition. It should be emphasized that this equilibrium

picked out the three regions where the hues of brighter and less bright spectra were identical. Exner's procedure is, of course, subject to important chromatic contrast effects, and does not lend itself to precise quantification.

TABLE I

SPECTRAL HUE-INVARIANT LOCI

Neutrally adapted right eye (10-min. preliminary dark adaptation), three *O*s. Entries show the average determination of each invariant locus in a single experimental session, and the root-mean-square deviation of the individual measurements (*SD*), with $N = 5$. The mean of the five average determinations (Grand Mean), the root-mean-square deviation about the Grand Mean (SD_M), and the mean of the five deviations (Mean Deviation) are also shown, with $N = 25$. All data are expressed in millimicrons.

	Locus of Blue Invariant		Locus of Green Invariant		Locus of Yellow Invariant	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Observer H	463.6	1.85	487.2	0.57	585.7	1.92
	470.7	1.64	492.5	1.87	584.0	1.37
	465.2	2.39	491.5	3.28	586.3	1.56
	467.6	0.96	493.1	3.68	591.6	2.61
	463.6	2.61	489.6	0.96	587.2	1.68
Grand Mean	466.1		490.8		587.0	
SD_M	2.82		2.60		2.84	
Mean Deviation		1.89		2.07		1.83
Observer J	474.4	1.24	503.9	0.92	582.2	0.76
	478.5	1.45	500.8	0.87	581.3	1.29
	476.0	2.14	497.9	1.50	579.0	1.15
	475.8	1.28	496.6	1.16	577.5	1.25
	476.9	2.33	497.9	1.18	579.9	1.73
Grand Mean	476.3		499.4		580.0	
SD_M	1.51		2.94		1.86	
Mean Deviation		1.69		1.13		1.24
Observer B	476.6	1.09	495.9	0.55	582.2	0.84
	474.0	1.17	493.2	2.02	582.2	1.53
	478.8	1.61	491.9	1.52	580.4	1.50
	480.0	1.41	491.5	1.66	580.8	0.91
	482.3	1.19	496.3	1.48	584.1	2.01
Grand Mean	478.3		493.8		581.9	
SD_M	3.18		2.23		1.46	
Mean Deviation		1.29		1.55		1.36

TABLE II

SPECTRAL HUE-INVARIANT LOCI WITH ADAPTATION AS VARIANT

Monocular observations with the right eye for three *O*s. The color temperature of the adapting field is expressed in degrees Kelvin. Size of adapting field is 46.8°; adaptation luminance is 25 millilamberts; and adaptation time is 5 min. The average determination of each invariant locus, the root-mean-square deviation of the individual measurements (*SD*), and the sum of the displacements ($\Sigma\Delta$'s) of the adapted invariant loci from the corresponding loci for the neutral state are all expressed in millimicrons. $N = 5$ except that $N = 25$ for the average determinations for the neutral state of adaptation.

Observer	Adaptation	Locus of Blue Invariant		Locus of Green Invariant		Locus of Yellow Invariant		$\Sigma\Delta$'s
		Mean	*SD*	Mean	*SD*	Mean	*SD*	
H	Neutral	466.1	1.89	490.8	2.07	587.0	1.83	
	10,000° K	466.6	1.56	498.5	1.17	607.1	3.32	28.3
	7500° K	471.3	1.64	504.9	1.34	585.5	2.89	20.8
	5500° K	467.0	2.21	495.4	0.82	589.7	3.46	8.2
	4800° K	478.1	1.34	505.2	2.36	605.8	2.77	45.2
	4300° K	463.9	2.13	505.2	1.04	605.5	2.64	35.1
J	Neutral	476.3	1.69	499.4	1.13	580.0	1.24	
	10,000° K	473.1	1.17	497.1	1.87	587.1	1.23	12.6
	7500° K	476.5	1.77	497.7	1.04	576.4	0.49	5.5
	5500° K	475.9	2.32	497.6	2.35	577.8	0.28	4.4
	4800° K	476.2	1.08	496.2	1.53	581.8	1.14	5.1
	4300° K	480.6	3.10	501.5	1.05	580.6	1.06	7.0
B	Neutral	478.3	1.29	493.8	1.55	581.9	1.36	
	10,000° K	472.4	2.16	493.2	0.84	588.6	2.10	13.2
	7500° K	479.7	1.79	495.3	2.11	584.0	1.00	5.0
	5500° K	473.4	0.82	495.2	2.36	586.0	1.28	10.5
	4800° K	481.1	1.88	498.6	1.39	586.1	1.67	13.8
	4300° K	476.6	1.52	497.6	0.82	586.8	1.30	10.2

refers only to the balance of *chromatic sensitivities;* general light sensitivity, of course, undergoes significant increases during considerably longer periods of dark adaptation. Moreover, the 10-min. interval empirically determined in these experiments is adequate only for the given circumstances, and this interval would not necessarily be adequate following pre-exposure to either highly chromatic or very intense illuminations. The duration of the latter effects requires a complete and separate investigation.

In the experiment proper, a 10-min. period of preliminary dark adaptation was consequently employed to bring *O* to a neutral state. *O* then made a single determination of the wave length of one of the three invariant hues in the spectrum, using the bracketing technique described above. A maximum of 10 sec. was allowed for a single observation, and ordinarily a determination was made in considerably less time than this. Following each observation *O* was given a 2-min. rest period in the dark. During a single experimental session, *O* made five determinations of each of the three invariant loci, taken in random order. The experiment was repeated for the neutral state of adaptation on five different days.

For determinations of the invariant loci with the eye adapted to various illuminants, the 10-min. period of preliminary dark adaptation was followed by a 5-min. light adaptation to a given color temperature. Following each single determination of the wave length of an invariant hue, *O* again viewed the adapting field for a period of 2 min. before making the next invariant locus determination. Again, five determinations of each of the three invariant loci constituted a single experiment. All observations were monocular, each *O* using the right eye.

RESULTS

Results for three *O*s for the neutral state of adaptation are presented in Table I. The table shows the average wave length of the individual determinations ($N = 5$) of the three in-

variant loci obtained in the separate experiments, and the root-mean-square deviations of the individual measurements. The grand average ($N = 25$) of the mean determinations for the five different days, the average of the five deviations, and the root-mean-square deviations of the five mean loci are also shown. All data are expressed in millimicrons. Examination of the table shows that, although the wave lengths of all three spectral loci may differ significantly from one *O* to the next, the results for any given *O* are both very precise and highly consistent.

The results for the various adaptive states are shown in Table II. The table shows the average invariant locus in the blue, green, and yellow spectral regions, respectively, for the neutral state of adaptation and for the eye adapted to the five different color temperatures. The root-mean-square deviations of the individual measurements, and the shifts of the three invariant loci from the corresponding loci for the neutral state, summed without regard to sign, are shown under the appropriate column headings. All results are expressed in millimicrons.

The invariant loci for these different conditions of bright adaptation can be determined with about the same degree of precision as for the neutral state of adaptation. The average wave lengths of the three invariant hues, however, do not usually agree with the average wave lengths determined for the neutral condition. The amount and direction of the shifts in these average determinations vary with the color temperature of the adapting stimulus, the particular invariant locus in question, and with the individual *O*. The results are shown graphically in Fig. 1, where the color temperature of the adapting illuminant is plotted on the abscissa, and the sum of the displacements of the three invariant loci from the corresponding neutral loci, in millimicrons, is on the ordinate.

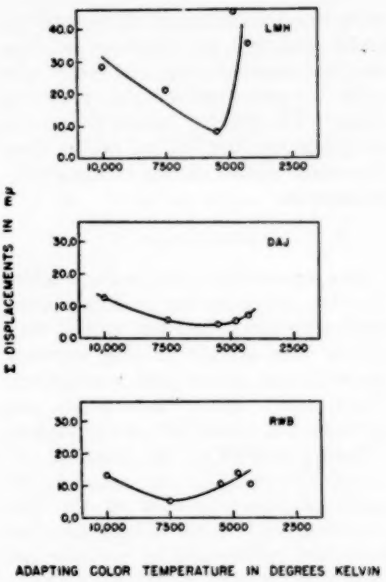

FIG. 1. Shifts in spectral hue-invariant loci with adaptation to various "whites"

The total amount of invariant hue displacement induced by adaptation to a given color temperature is quite different from one *O* to the next. In all cases, however, the function exhibits a minimum, and there is a systematic increase in the total displacement with increasing departures from the color temperature of the minimal region. For observers *H* and *J*, the minimal total displacement occurs for the adapting field of 5500° K, and the total shift increases for adaptations to color temperatures both higher and lower than this value. For the third observer, *B*, the minimal displacement occurs for the 7500° K adaptation. It should be noted that

more exact location of these minima could probably be achieved by employing smaller steps between the color temperatures of the adapting fields. The present values define the color temperature region, rather than the exact locus, of the minimal displacement.

Discussion

Our variability data make it clear that the shifts in the invariant hues cannot be accounted for by the day-to-day fluctuations in the intrinsic properties of the visual mechanism. These shifts must, accordingly, be attributed to the effects of the various adapting stimuli on the state of the visual system. If we assume the index chosen (the sum of the displacements of the three invariant loci from the corresponding loci for the neutral condition) to be valid, we have a quantitative expression for the degree of change in chromatic tuning brought about by the various adaptations. The questions posed at the outset of the experiment can now be answered.

All the color temperatures employed in this experiment appeared neutral at the end of the 5-min. adapting period. The physiological responses of the visual mechanism, however, were obviously not identical in all cases. Those color temperatures, moreover, whose durative action produces a minimal total displacement in the wave lengths of the spectral invariant hues can be considered to approximate most closely true neutral or white stimuli for our *O*s. These stimuli not only appeared neutral, they also produced a minimal change in the initially neutral state of the visual mechanism. For two *O*s, this color temperature is 5500° K, and for the third, 7500° K. The increasingly greater shifts which occurred for color temperatures which depart more and more from the "neutral" ones in either direction can be taken as evidence of progressively greater imbalances in the relative chromatic sensitivities. In these instances, "whites" which qualify on the basis of a perceptual criterion are clearly disqualified on the basis of the physiological index.

We have, in this discussion, deliberately emphasized the notion of a balance or imbalance of chromatic sensitivities and have avoided any attempt to account for the shifts in the spectral invariant loci in terms of inferred changes in specific spectral response curves. Certainly continued stimulation by a predominantly blue stimulus must desensitize the blue-sensitive receptor or neural elements, and the same kind of desensitization must occur for stimuli of other chromaticities. A detailed and meaningful analysis, however, could only be made on the basis of calculated changes in spectral response curves which are theoretically derived from additional independent experimental evidence. At the moment there are two major difficulties: neither the number nor the forms of the spectral response curves are agreed upon (**15**, pp. 54–57, **17**, pp. 368–379), and the relevant experimental data which are available are the averaged data of other *O*s, obtained under different circumstances, for different retinal areas, etc. In our opinion a few vague guesses about specific mechanisms would add nothing at the moment to the primary purpose of this study, which is to present an experimental technique for determining reliable white standards for visual experimentation. The use of such standards can aid in the accumulation of comparable experimental results and, in turn, may enable us eventually

to make more meaningful and fruitful inferences concerning the specific nature and behavior of the visual processes.

Summary

The use of white standards in visual experiments requires specification, not only of the physical properties of, and the perceptual response to, the standard stimulus, but also of the state of the visual system in relation to the continued action of such a white standard. A technique for determining the effects of the continued action of various perceptually white or neutral stimuli on the balance of visual sensitivities has been developed and described. The technique involves measurement of the locations of the spectral wave lengths which are perceived as invariant in hue with changes in intensity with the visual system in an initial neutral equilibrium condition, and redetermination of these loci following the durative action of a series of color temperatures all of which evoke a perceptually neutral response. The total displacements in the three spectral invariant loci following exposure to the various color temperatures are found to pass through a minimum at 5500° K for two *O*s and at 7500° K for a third. Color temperatures which depart more and more from the regions of the minima induce progressively larger shifts in the wave lengths of the spectral invariant hues. It is assumed that the color temperatures which produce minimal displacements most closely approximate true neutral stimuli for these *O*s, and that the shifts induced by the other color temperatures are evidence of progressively greater amounts of chromatic imbalance in the visual system.

(Manuscript received for immediate publication May 7, 1951)

References

1. Cohen, J. Color adaptation of the human eye. *Amer. J. Psychol.*, 1946, **59**, 84–110.
2. Dittler, R., and Satake, Y. Eine Methode zur Bestimmung der gegenfarbig wirkenden Wellenlangen des Spectrums. *Z. Sinnesphysiol.*, 1914, **48**, 240–251.
3. Dreher, E. Methodische Untersuchung der Farbentonänderungen homogener Lichter bei zunehmend indirektem Sehen und veränderter Intensität. *Z. Sinnesphysiol.*, 1912, **46**, 1–82.
4. Exner, F. Ueber die Grundempfindungen im Young-Helmholtzschen Farbensystem. *Sitzber. Akad. Wiss., Wien, Math.-naturw. Kl.*, 1902 (Abt. 2a), **111**, 857–877.
5. Forbes, W. T. M. What is white? *Amer. J. Psychol.*, 1928, **40**, 522–523.
6. Hering, E. Ueber individuelle Verschiedenheiten des Farbensinnes. *Lotos, Jahr. Naturwiss. N. F.*, 1885, **6**, 142–198.
7. Hunt, R. W. G. Visual adaptation and the apparent saturation of colours. *Proc. phys. Soc. Lond.*, 1949, **62**, 203–206.
8. Hurvich, L. M., and Jameson, Dorothea. A psychophysical study of white. I. Neutral adaptation. *J. opt. Soc. Amer.* 1951, **41**, 521–527.
9. Jameson, Dorothea, and Hurvich, L. M. A psychophysical study of white. II. Neutral adaptation. Area and duration as variants. *J. opt. Soc. Amer.* 1951, **41**, 528–536.
10. Purdy, D. M. Chroma as a function of retinal illumination. Unpublished Doctor's dissertation, Harvard Univ., 1929.
11. Purdy, D. M. Spectral hue as a function of intensity. *Amer. J. Psychol.*, 1931, **43**, 541–559.
12. Troland, L. T. The colors produced by equilibrium photopic adaptation. *J. exp. Psychol.*, 1921, **4**, 344–390.
13. Tschermak, A. Was ist neutrales Licht? *Arch. Augenheilk.*, 1930, **103**, 337–356.
14. Tschermak, A. Licht- und Farbensinn. In A. Bethe *et al.* (Ed.), *Handb. norm. path. Physiol.* Berlin: Springer, 1929, **12**, (1), 295–501.
15. Tschermak-Seysenegg, A. v. *Einführung in die physiologische Optik.* (2nd Ed.) Wien: Springer, 1947.
16. Wright, W. D. The measurement and analysis of colour adaptation phenomena. *Proc. roy. Soc.*, Ser. B, 1934, **115**, 49–87.
17. Wright, W. D. *Researches in normal and defective colour vision.* St. Louis: C. V. Mosby, 1947.

Classics Among

PSYCHOLOGICAL MONOGRAPHS

Thorndike, E. L. The Mental Life of the Monkey. 1899, #15. *$.50*

Carr, Harvey. Visual Illusion of Movement During Eye Closure. 1905, #31. *$1.25*

Watson, John B. Kinaesthetic and Organic Sensations: Their Role in the Reactions of the White Rat to the Maze. 1907, #33, *$1.00*

Shepherd, William T. Some Mental Processes of the Rhesus Monkey. 1910, #52. *$.75*

Franz, Shepard Ivory and Lafora, Gonzalo R. On the Functions of the Cerebrum: The Occipital Lobes. 1911, #56. *$1.25*

Fernberger, Samuel W. On the Relation of the Methods of Just Perceptible Differences and Constant Stimuli. 1912, #61. *$1.00*

Boring, Edwin G. Learning in Dementia Praecox. 1913, #63. *$1.00*

Langfeld, Herbert S. On the Psychophysiology of a Prolonged Fast. 1914, #71. *$.75*

Franz, Shepard Ivory. I. Symptomological Differences Associated with Similar Cerebral Lesions in the Insane. II. Variations in Distribution of the Motor Centers. 1915, #81. *$1.50*

Peckstein, Louis Augustus. Whole versus Part Methods in Motor Learning. 1917, #99. *$.75*

Kjerstad, Conrad L. The Form of the Learning Curves for Memory. 1919, #116. *$1.35*

Tolman, Edward C. Retroactive Inhibition as Affected by Conditions of Learning. 1918, #107. *$.75*

MANY OF THE EARLY MONOGRAPHS ARE OUT-OF-PRINT. ONLY A LIMITED QUANTITY OF THE ABOVE NUMBERS ARE AVAILABLE.

AMERICAN PSYCHOLOGICAL ASSOCIATION

1515 Massachusetts Ave. N.W., Washington 5, D. C.

American Psychological Association

1515 Massachusetts Ave. N.W.
Washington 5, D. C.

Publications:

AMERICAN PSYCHOLOGIST

Editor: Fillmore H. Sanford, *American Psychological Association.* Contains all official papers of the Association and articles concerning psychology as a profession; monthly.

Subscription: $7.00 (Foreign $7.50). Single copies, $.75.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

Editor: J. McV. Hunt, *Institute of Welfare Research, New York City.* Contains original contributions in the field of abnormal and social psychology, reviews, and case reports; quarterly.

Subscription: $6.00 (Foreign $6.50). Single copies, $1.75.

JOURNAL OF APPLIED PSYCHOLOGY

Editor: Donald G. Paterson, *University of Minnesota.* Contains material covering applications of psychology to business, industry, and education; bi-monthly.

Subscription: $6.00 (Foreign $6.50). Single copies, $1.25.

JOURNAL OF COMPARATIVE AND PHYSIOLOGICAL PSYCHOLOGY

Editor: Harry F. Harlow, *Department of the Army, Washington, D. C.* Contains original contributions in the field of comparative and physiological psychology; bi-monthly.

Subscription: $7.00 (Foreign $7.50). Single copies, $1.25.

JOURNAL OF CONSULTING PSYCHOLOGY

Editor: Laurance F. Shaffer, *Teachers College, Columbia University.* Contains articles in the field of clinical and consulting psychology, counseling and guidance; bi-monthly.

Subscription: $5.00 (Foreign $5.50). Single copies, $1.00.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Editor: Arthur W. Melton, *U. S. Air Force, San Antonio, Texas.* Contains original contributions of an experimental character; monthly, two volumes per year.

Subscription: $14.00 (Foreign $14.50). Single copies, $1.25.

PSYCHOLOGICAL ABSTRACTS

Editor: C. M. Louttit, *University of Illinois.* Contains noncritical abstracts of the world's literature in psychology and related subjects; monthly.

Subscription: $7.00 (Foreign $7.50). Single copies, $.75.

PSYCHOLOGICAL BULLETIN

Editor: Lyle H. Lanier, *University of Illinois.* Contains critical reviews of psychological literature, methodological articles, book reviews, and discussions of controversial issues; bi-monthly.

Subscription: $7.00 (Foreign $7.50). Single copies, $1.25.

PSYCHOLOGICAL MONOGRAPHS: GENERAL AND APPLIED

Editor: Herbert S. Conrad, *U. S. Office of Education, Washington, D. C.* Contains longer researches and laboratory studies which appear as units; published at irregular intervals, about ten numbers per year.

Subscription: $6.00 per volume (Foreign $6.50). Single copies, price varies according to size.

PSYCHOLOGICAL REVIEW

Editor: Carroll C. Pratt, *Princeton University.* Contains original contributions of a theoretical nature; bi-monthly.

Subscription: $5.50 (Foreign $6.00). Single copies, $1.00.